

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 17:15:11 ; Search time 47 Seconds
(without alignments)
1867.570 Million cell updates/sec

Title: US-09-593-793a-113
Perfect score: 2861
Sequence: 1 MVQRLVSVSLRLRRKQAQLL.....AIYFATQVVDKSLAKYSA 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_19Jun03.*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2861	100.0	553	19 AAW71869	Amino acid encoded
2	2861	100.0	553	19 AAW69385	Prostate tumour sp
3	2861	100.0	553	21 AAB28527	Protein encoded by
4	2861	100.0	553	21 AAY82002	Human immunogenic
5	2861	100.0	553	22 AAU69763	Human prostate cDN
6	2861	100.0	553	22 AAU04961	Human prostate tum
7	2861	100.0	553	22 AAU01117	Human prostate-spe
8	2861	100.0	553	22 AAG99002	Human prostate-spe
9	2861	100.0	553	22 AAG62150	Human P501S invent

10	2861	100.0	553	22 AAB74800	Prostate tumour an
11	2861	100.0	553	23 ABG94411	Human prostate tum
12	2861	100.0	553	23 ABG76665	Prostate tumour pr
13	2861	100.0	553	23 ABG77575	Human mast cell re
14	2861	100.0	553	23 ABG61900	Prostate cancer-as
15	2861	100.0	553	23 ABG52222	Human L1-12 protei
16	2861	100.0	553	23 AAU10324	Human PROST 03. H
17	2861	100.0	553	23 AAU82643	Human breast tumou
18	2861	100.0	553	24 ABU71653	Prostate cancer sp
19	2861	100.0	1079	22 ABU74830	Prostate tumour an
20	2861	100.0	1079	22 ABU71860	Prostate specific
21	2601	90.9	710	23 AAU50661	Thioredoxin-ubiqui
22	2596	90.7	595	22 AAU01318	Alpha prepro-P501S
23	1696	59.3	359	24 ABU71887	Human prostate spe
24	1677.5	58.6	530	23 AAU50662	Thioredoxin-ubiqui
25	1417.5	49.5	371	22 AAU69875	Human prostate cDN
26	1417.5	49.5	371	22 AAU01230	P553S splice varia
27	1417.5	49.5	371	23 ABG53335	Human P553S splice
28	1417.5	49.5	371	24 ABU71766	Prostate cancer as
29	1416	49.5	371	22 AAU01362	Human gene 11 enco
30	1403.5	49.5	371	23 ABG64105	Human albumin fusi
31	1403.5	49.1	400	22 AAU69907	Human prostate pro
32	1403.5	49.1	400	22 AAU01262	Ra12-P501S-E2 cons
33	1403.5	49.1	400	23 ABG95367	Ra12-P501S-E2 cons
34	1403.5	49.1	400	24 ABU71798	Prostate cancer as
35	1287	45.0	255	20 AAW85068	Protein encoded by
36	1287	45.0	255	21 AAB29268	Human prostate-rel
37	1287	45.0	255	22 AAU04205	Prostate-specific
38	1187	45.0	255	23 AAU019084	Human prostate-spe
39	1150	40.2	231	21 AAY54369	Amino acid sequenc
40	1120	39.1	252	22 AAU01423	Human secreted pro
41	1014.5	35.5	326	23 ABG77571	Human mast cell re
42	519.5	18.2	599	22 ABG60709	Drosophila melanog
43	457.5	16.0	748	22 AAU40227	Human polypeptide
44	426	14.9	123	22 AAU69873	Human prostate cDN
45	426	14.9	123	22 AAU01228	P553S splice varia

ALIGNMENTS

RESULT 1
AAW71869
ID AAW71869 standard; Protein; 553 AA.
XX
AC AAW71869;
XX
XX
DT 06-JAN-1999 (first entry)
XX
DE Amino acid encoded by prostate tumour clone L1-12.
XX
KW Prostate; cancer; tumour; vaccine; immunogen; clone.
XX
OS Homo sapiens.
XX
PN WO9837093-A2.
XX
PD 27-AUG-1998.
XX
PF 25-FEB-1998; 98WO-US03492.
XX
PR 09-FEB-1998; 98US-0020956.
PR 25-FEB-1997; 97US-0806099.
PR 01-AUG-1997; 97US-0904804.
XX
PA (CORI-) CORIXA CORP.
XX
PI Dillon DC, Xu J;
XX
DR WPI; 1998-609886/51.
XX
XX N-PSDB; AAV61201.
XX
PT Polypeptides comprising immunogenic portions of prostate proteins -

used in a vaccine for the treatment of prostate cancer

Example 1; Page 82-84; 130pp; English.

The present sequence is an immunogenic portion of a prostate tumour protein. The immunogen, or the DNA encoding it, can be used as a vaccine for the treatment of prostate cancer. The immunogen was isolated from a prostate tumour cDNA library obtained by subtracting a prostate tumour cDNA expression library with a normal tissue cDNA library.

Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 19; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRRHKAQLLVNLLTFGLEVCLAAGITYVPPLLEVGVEEKFMTVLGIG 60
DB 1 MVQRLWVSRLRRHKAQLLVNLLTFGLEVCLAAGITYVPPLLEVGVEEKFMTVLGIG 60
QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
QY 121 ELALLILGVGLDFCGOVCFPLEALLSDFRDPHCRQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGOVCFPLEALLSDFRDPHCRQAYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTSAALPYLGTQECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
DB 181 IDWDTSAALPYLGTQECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
QY 241 CCPCARLAFRNLGALLPRHLQCCMRPTLRRLFVAELCSWMLMTFTFLFYDFVGEGL 300
DB 241 CCPCARLAFRNLGALLPRHLQCCMRPTLRRLFVAELCSWMLMTFTFLFYDFVGEGL 300
QY 301 YQGVPRABPTGTEARRHYDEGVRMGSLGLFQCAISLVFSLVMDRLVQRFTRAVYLASVA 360
DB 301 YQGVPRABPTGTEARRHYDEGVRMGSLGLFQCAISLVFSLVMDRLVQRFTRAVYLASVA 360
QY 361 AFPVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 AFPVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGSPTEA 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGSPTEA 480
QY 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYPATQ 540
DB 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYPATQ 540

RESULT 2

AAW69385

ID AAW69385 standard; Protein; 553 AA.

XX

AC AAW69385;

XX

DT 25-MAR-2003 (updated)

DT 08-DEC-1998 (first entry)

XX

DE Prostate tumour specific gene clone L1-12 protein.

XX

KW Prostate tumour specific gene; human; prostate cancer; detection;

KW therapy.

XX

OS Homo sapiens.

WO9837418-A2.

XX

PN

XX

PD

XX

25-FEB-1998; 98WO-US03690.

XX

XX

25-FEB-1997; 97US-0806596.

XX

PR

01-AUG-1997; 97US-0904809.

XX

PR

09-FEB-1998; 98US-0020747.

XX

(CORI-) CORIXA CORP.

XX

PI

Dillon DC, Xu J;

XX

WPI; 1998-480805/41.

DR

N-PSDB; AAV58586.

DR

XX

PT

Novel human prostate specific tumour protein and fragments - useful

for detecting and treating prostate cancers

XX

PS

Example 1; Page 87-89; 141pp; English.

XX

CC

This sequence is encoded by a human prostate tumour specific gene, and

can be used in the method of the invention. The method is for detecting

prostate cancer comprises contacting a biological sample with an agent

able to bind an immunogenic portion of a prostate protein (such as

this protein sequence). An antibody which binds to an immunogenic

portion of the prostate protein, and the method can be used to detect,

monitor progression of, or treat prostate cancers. The antibody may

also be conjugated to a therapeutic agent for use in therapy of prostate

cancers.

CC

(Updated on 25-MAR-2003 to correct PR field.)

CC

XX

SQ

Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 19; Length 553;

Best Local Similarity 100.0%; Pred. No. 2.2e-269;

Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRRHKAQLLVNLLTFGLEVCLAAGITYVPPLLEVGVEEKFMTVLGIG 60

DB 1 MVQRLWVSRLRRHKAQLLVNLLTFGLEVCLAAGITYVPPLLEVGVEEKFMTVLGIG 60

QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120

DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120

QY 121 ELALLILGVGLDFCGOVCFPLEALLSDFRDPHCRQAYSVYAFMISLGGCLGYLLPA 180

DB 121 ELALLILGVGLDFCGOVCFPLEALLSDFRDPHCRQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTSAALPYLGTQECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

DB 181 IDWDTSAALPYLGTQECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY 241 CCPCARLAFRNLGALLPRHLQCCMRPTLRRLFVAELCSWMLMTFTFLFYDFVGEGL 300

DB 241 CCPCARLAFRNLGALLPRHLQCCMRPTLRRLFVAELCSWMLMTFTFLFYDFVGEGL 300

QY 301 YQGVPRABPTGTEARRHYDEGVRMGSLGLFQCAISLVFSLVMDRLVQRFTRAVYLASVA 360

DB 301 YQGVPRABPTGTEARRHYDEGVRMGSLGLFQCAISLVFSLVMDRLVQRFTRAVYLASVA 360

QY 361 AFPVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

DB 361 AFPVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGSPTEA 480

DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGSPTEA 480

QY 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYPATQ 540

Db 481 RVVPGRGICLDLAIDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 3
AAB28527
ID AAB28527 standard; Protein; 553 AA.
AC AAB28527;
XX
DT 07-FEB-2001 (first entry)
XX
DE Protein encoded by human breast tumour cDNA clone P501S.
XX
KW Human; breast tumour antigen; cytostatic; immunotherapy;
KW breast cancer; vaccine.
XX
OS Homo sapiens.
XX
PN WO200061756-A2.
XX
PD 19-OCT-2000.
XX
PF 10-APR-2000; 2000WO-US09688.
XX
PR 09-APR-1999; 99US-0288950.
PR 02-JUL-1999; 99US-0346327.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Xu J, Dillon DC;
XX
DR WPI; 2000-638568/61.
DR N-PSDB; AAC79473.
XX
PT A novel isolated polypeptide comprising an immunogenic portion of a
PT breast cancer protein useful in the detection and treatment of breast
PT cancer -
XX
PS Claim 2; Page 92-93; 95pp; English.
XX
CC The present sequence is encoded by a cDNA sequence which was isolated
CC from a breast tumour cDNA library. It is provided in a specification
CC relating to compounds for immunotherapy and diagnosis of breast cancer.
CC Breast tumour antigens and the polynucleotides that encode them may be
CC used in the production of a pharmaceutical composition to be used in the
CC treatment of breast cancer. Proliferated T cells and incubated antigen
CC presenting cells are also required. The polypeptides and polynucleotides
CC may also be used to produce a vaccine.
XX
SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 21; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLLRRHKAQLLLVNLITFGLEVCLAAGITYVPPLLELVGVVEEKFMTWVLGIG 60
Db 1 MVQRLWVSRLLRRHKAQLLLVNLITFGLEVCLAAGITYVPPLLELVGVVEEKFMTWVLGIG 60
Qy 61 PVLGLVCVPLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
Db 61 PVLGLVCVPLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
Qy 121 ELALLILGVLLDFCCQVCTPPEALLSDFRDPDHCROAYSFYAFWISLGGCLGYLLPA 180
Db 121 ELALLILGVLLDFCCQVCTPPEALLSDFRDPDHCROAYSFYAFWISLGGCLGYLLPA 180
Qy 181 IDWDTSAAPYLGTQBECLFGLLTFLTCVATLLVAEEAALGPTPEAGLSAPSLSPH 240

Db 181 IDWDTSAAPYLGTQBECLFGLLTFLTCVATLLVAEEAALGPTPEAGLSAPSLSPH 240
Qy 241 CCPCRLARLAFRNIGALLPELHOLCCRMPTLRRLFVAELCSWMLMTFTFLFYDFVGEGL 300
Db 241 CCPCRLARLAFRNIGALLPELHOLCCRMPTLRRLFVAELCSWMLMTFTFLFYDFVGEGL 300
Qy 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFGTTRAVYLASVA 360
Db 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFGTTRAVYLASVA 360
Qy 361 APFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKPGAPPFNGHVAGGSGLLPPPPALCGASACDVSRRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPPFNGHVAGGSGLLPPPPALCGASACDVSRRVVVGEPTEA 480
Qy 481 RVVPGRGICLDLAIDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
Db 481 RVVPGRGICLDLAIDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 4
AAY82002
ID AAY82002 standard; Protein; 553 AA.
XX
AC AAY82002;
XX
DT 13-JUN-2000 (first entry)
XX
DE Human immunogenic prostate tumour protein sequence SEQ ID NO:113.
XX
KW Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW immunogenic; cytostatic; vaccine.
XX
OS Homo sapiens.
XX
PN WO200004149-A2.
XX
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-US15838.
XX
PR 14-JUL-1998; 98US-0115453.
PR 23-SEP-1998; 98US-0116134.
PR 23-SEP-1998; 98US-0159812.
PR 15-JAN-1999; 99US-0159822.
PR 15-JAN-1999; 99US-0232149.
PR 09-APR-1999; 99US-0232880.
XX
XX (CORI-) CORIXA CORP.
PA
PI Dillon DC, Harlocker SL, Yugu J, Xu J, Mitcham JL;
XX
DR WPI; 2000-171268/15.
XX
PT New polypeptide useful for treating and diagnosing prostate cancer
PT comprises an immunogenic portion of prostate tumor protein -
XX
PS Claim 3; Page 138-139; 263pp; English.
XX
CC The present invention describes isolated polypeptides, comprising an
CC immunogenic portion of a prostate tumor protein (PTP). The polypeptides
CC and polynucleotides encoding them have cytostatic activity and can be
CC used in vaccines and in gene therapy. The polypeptides and
CC polynucleotides encoding them, antigen presenting cells which express

CC the polypeptides, antibodies against the polypeptides and vaccines
 CC comprising them can be used for inhibiting the development of prostate
 CC cancer in a patient. The polypeptides can be used to generate antibodies
 CC or anti-idiotypic antibodies for passive immuno therapy. A portion of
 CC the polynucleotides encoding the polypeptides can be used as a probe or
 CC to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
 CC AA82000 to AA82020 represent sequences used in the exemplification of
 CC the present invention.

XX SQ Sequence 553 AA;
 Query Match 100.0%; Score 2861; DB 21; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLEVGVEKFWMTVLGIG 60
 DB 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLEVGVEKFWMTVLGIG 60
 QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
 DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
 QY 121 ELALLILGVGLLDFCGQVCFTEALLSDLPDPHCRQAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVGLLDFCGQVCFTEALLSDLPDPHCRQAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDWTSALAPYLGTQECLFGLLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 DB 181 IDWTSALAPYLGTQECLFGLLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 QY 241 CCPCRLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWALMTFTLYTDFVGEGL 300
 DB 241 CCPCRLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWALMTFTLYTDFVGEGL 300
 QY 301 YQGVPRAEPTGEARRHYDEGVMSGLGLFLOCAISLVFSLVMDRLVQRTGTRAVYLASVA 360
 DB 301 YQGVPRAEPTGEARRHYDEGVMSGLGLFLOCAISLVFSLVMDRLVQRTGTRAVYLASVA 360
 QY 361 AFPVAAGATCLSHSVAVVTASAALTGFTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 AFPVAAGATCLSHSVAVVTASAALTGFTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY 421 ASSEDSLMTSFLPGPKGAPFNGHVAGAGSGLLPPPPALCGASACDVSVRVVVGTEA 480
 DB 421 ASSEDSLMTSFLPGPKGAPFNGHVAGAGSGLLPPPPALCGASACDVSVRVVVGTEA 480
 QY 481 RVVPGRGICLDLALDSAFLLSOVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
 DB 481 RVVPGRGICLDLALDSAFLLSOVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
 QY 541 VVFDKSLAKYSA 553
 DB 541 VVFDKSLAKYSA 553

RESULT 5
 AAU69763
 ID AAU69763 standard; Protein; 553 AA.
 XX
 XX AAU69763;
 XX
 XX 30-JAN-2002 (first entry)
 XX
 XX Human prostate cDNA encoded protein #3.
 XX Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
 XX
 XX Homo sapiens.
 XX OS
 XX PN W0200173032-A2.
 XX
 XX 04-OCT-2001.

XX PF 27-MAR-2001; 2001WO-US09919.
 XX PR 27-MAR-2000; 2000US-0536857.
 PR 09-MAY-2000; 2000US-0568100.
 PR 12-MAY-2000; 2000US-0570737.
 PR 13-JUN-2000; 2000US-0593793.
 PR 17-JUN-2000; 2000US-0605783.
 PR 10-AUG-2000; 2000US-0636215.
 PR 29-AUG-2000; 2000US-0651236.
 PR 06-SEP-2000; 2000US-0657279.
 PR 02-OCT-2000; 2000US-0679426.
 PR 10-OCT-2000; 2000US-0685166.
 XX (CORI-) CORIXA CORP.
 PA Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedwick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
 XX WPI; 2001-639232/73.
 DR N-PSDB; AAS63557.
 XX
 PT New human prostate-specific polypeptides and polynucleotides useful for
 the diagnosis and treatment of cancer, especially prostate cancer -
 XX Claim 2; Page 269-270; 579pp; English.
 CC The invention relates to isolated prostate-specific
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
 CC antibodies raised against the polypeptides for antigenic epitopes
 CC derived from them) and antigen-presenting cells expressing the
 CC polypeptides. The antibodies are useful for detecting the presence of
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
 CC the antigen-presenting cells are useful for stimulating and/or expanding
 CC T cells specific for a tumour protein, and for inhibiting the development
 CC of cancer especially prostate cancer. Compositions comprising the
 CC polynucleotide and/or polypeptide are useful for stimulating an immune
 CC response, and for treating cancer. The oligonucleotide is useful for
 CC detecting cancer. The present sequence is a prostate specific
 CC polypeptide of the invention.
 XX SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 22; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLEVGVEKFWMTVLGIG 60
 DB 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLEVGVEKFWMTVLGIG 60
 QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
 DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
 QY 121 ELALLILGVGLLDFCGQVCFTEALLSDLPDPHCRQAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVGLLDFCGQVCFTEALLSDLPDPHCRQAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDWTSALAPYLGTQECLFGLLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 DB 181 IDWTSALAPYLGTQECLFGLLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 QY 241 CCPCRLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWALMTFTLYTDFVGEGL 300
 DB 241 CCPCRLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWALMTFTLYTDFVGEGL 300
 QY 301 YQGVPRAEPTGEARRHYDEGVMSGLGLFLOCAISLVFSLVMDRLVQRTGTRAVYLASVA 360
 DB 301 YQGVPRAEPTGEARRHYDEGVMSGLGLFLOCAISLVFSLVMDRLVQRTGTRAVYLASVA 360
 QY 361 AFPVAAGATCLSHSVAVVTASAALTGFTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420


```
Db 361 APPVAAGATCLSHSAVAVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIFYATQ 540
QY 541 VVFDKSDIAKYSA 553
Db 541 VVFDKSDIAKYSA 553

RESULT 6
AAU04961
ID AAU04961 standard; Protein; 553 AA.
AC AAU04961;
DT 24-OCT-2001 (first entry)
DE Human prostate tumour protein L1-12.
KW Human; prostate tumour protein; prostate cancer.
XX Homo sapiens.
OS Homo sapiens.
PN US6262245-B1.
PD 17-JUL-2001.
PF 25-FEB-1998; 98US-0030607.
PR 25-FEB-1997; 97US-0806099.
PR 01-AUG-1997; 97US-0904804.
PR 09-FEB-1998; 98US-0020956.
XX (CORI-) CORIXA CORP.
XX Xu J, Dillon DC;
PI WPI; 2001-440862/47.
DR N-PSDB; AAS10108.
XX Novel polynucleotide encoding polypeptide comprising a portion of
PT prostate tumour protein useful for inhibiting development of prostate
PT cancer or for treating prostate cancer in a patient
XX Example 1; Column 125-127; 105pp; English.
XX The sequence is a partial prostate tumour protein, encoded by a prostate
CC tumour specific cDNA. The DNA is useful for inhibiting the development
CC of prostate cancer or for treating prostate cancer in a patient.
XX Sequence 553 AA;
SQ Query Match 100.0%; Score 2861; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLHRKQAQLLVNLLTFGLVCLAGITVPPPLLEVGVEEKFTMWLVGIG 60
Db 1 MVQRLWVSRLLHRKQAQLLVNLLTFGLVCLAGITVPPPLLEVGVEEKFTMWLVGIG 60
QY 61 PVGLVCPVLLGSADHWGRGRRPFIMWLSGILLSLFLIPAGWLAGLCPDPRPL 120
Db 61 PVGLVCPVLLGSADHWGRGRRPFIMWLSGILLSLFLIPAGWLAGLCPDPRPL 120
QY 121 ELALLILGVGLDFCGQVCFTPLEALLSDLFRDPDHCQAYSVYAFMISLGCLGYLLPA 180
```

```
Db 121 ELALLILGVGLDFCGQVCFTPLEALLSDLFRDPDHCQAYSVYAFMISLGCLGYLLPA 180
QY 181 IDWDTSAALPYLGTQEBCLFGLLTLLIFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
Db 181 IDWDTSAALPYLGTQEBCLFGLLTLLIFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
QY 241 CCPCARLAFRNLGALLPRLHQLCCMRPRTLRLRLEVAELCSMMALMTFTLFTYDFVGEGL 300
Db 241 CCPCARLAFRNLGALLPRLHQLCCMRPRTLRLRLEVAELCSMMALMTFTLFTYDFVGEGL 300
QY 301 YQGVPRAPGTEARHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRPRTGTRAVLASVA 360
Db 301 YQGVPRAPGTEARHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRPRTGTRAVLASVA 360
QY 361 APPVAAGATCLSHSAVAVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APPVAAGATCLSHSAVAVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIFYATQ 540
QY 541 VVFDKSDIAKYSA 553
Db 541 VVFDKSDIAKYSA 553

RESULT 7
AAU01117
ID AAU01117 standard; Protein; 553 AA.
AC AAU01117;
DT 04-OCT-2001 (first entry)
DE Human prostate-specific amino acid sequence L1-12.
KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW cytostatic; gene therapy; metastasis.
OS Homo sapiens.
XX WO200151633-A2.
PN 19-JUL-2001.
PD 16-JAN-2001; 2001WO-US01574.
PF 14-JAN-2000; 2000US-0483672.
PA (CORI-) CORIXA CORP.
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
PI Wang A, Meagher MJ;
XX WPI; 2001-425873/45.
XX New polynucleotide encoding a prostate-specific protein, for
PT diagnosing, monitoring and treating prostate cancer in a patient and
PT for use in vaccines -
XX Claim 2; Page 267-268; 543pp; English.
XX The present invention describes polynucleotide sequences (I) which encode
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II),
CC antibodies to (II), fusion proteins comprising (II), and isolated
CC T cells prepared using (I) or (II) are used treat cancer in a patient.
```

CC (I) and the antibodies are also used in the detection of cancer in a
 CC patient. The cancer that is diagnosed or treated is particularly
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
 CC (I) can be used for monitoring the progression of cancer in a patient.
 CC (I) and (II) can also be used to improve diagnostic and therapeutic
 CC methods for prostate cancer. They can indicate the level of metastasis
 CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to
 CC AAH01318 represent polynucleotide and amino acid sequences used in the
 CC exemplification of the present invention.

XX
 SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 22; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2,2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGTTPVPLLLVGVGVEKFTMTVLGIG 60
 DB 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGTTPVPLLLVGVGVEKFTMTVLGIG 60

QY 61 PVILGLVCPVLLGSASDHWRGRRPPTIWAISLIGLISLFLIPRAGWLAGLLCPDPRLL 120
 DB 61 PVILGLVCPVLLGSASDHWRGRRPPTIWAISLIGLISLFLIPRAGWLAGLLCPDPRLL 120

QY 121 ELALLILGVGLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVGLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTSAALPYLGTQBECLFGLLTILFTCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 DB 181 IDWDTSAALPYLGTQBECLFGLLTILFTCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY 241 CCPCARLAFRNLGALLPRHLCCRMPTLRLLFVAELCSWMLMTFTLFTYDFVGEGL 300
 DB 241 CCPCARLAFRNLGALLPRHLCCRMPTLRLLFVAELCSWMLMTFTLFTYDFVGEGL 300

QY 301 YQGVPRAEPTGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
 DB 301 YQGVPRAEPTGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRFQTRAVYLASVA 360

QY 361 APFVAAGATCLSHSVAVVTASAALTGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 APFVAAGATCLSHSVAVVTASAALTGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 DB 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIYFATQ 540
 DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIYFATQ 540

QY 541 VVFDKSDLAKYSA 553
 DB 541 VVFDKSDLAKYSA 553

RESULT 8
 AAG99002
 ID AAG99002 standard; Protein; 553 AA.
 XX
 AC AAG99002;
 XX
 DT 25-SEP-2001 (first entry)
 XX
 DE Human prostate-specific amino acid sequence L1-12/P501S.
 XX
 KW Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
 KW chromosome 22q11.2; prostate-specific protein; chromosome 1;
 KW prostate specific antigen; PSA.
 XX
 OS Homo sapiens.
 XX

PN WO200134802-A2.
 XX
 PD 17-MAY-2001.
 XX
 PF 09-NOV-2000; 2000WO-US30904.
 XX
 PR 12-NOV-1999; 99US-0439313.
 PR 18-NOV-1999; 99US-0443686.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kalos WD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;
 XX
 DR WPI; 2001-308785/32.
 XX
 PT Isolated polypeptide comprising at least an immunogenic portion of a
 PT prostate-specific protein, useful in the diagnosis and therapy of
 PT prostate cancer -
 XX
 PS Claim 3; Page 167-168; 325pp; English.
 XX
 CC The present invention describes an isolated polypeptide (P1) comprising
 CC at least an immunogenic portion of a prostate-specific protein, or its
 CC variant. Also described are polynucleotides (N1) encoding (P1) and
 CC (N1) have cytostatic activity and can be used in vaccine production.
 CC The polypeptides, nucleic acids and antibodies from the present
 CC invention are useful in the diagnosis and therapy of prostate cancer.
 CC Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
 CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
 CC region. Prostate specific antigen (PSA) P501S was located on
 CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
 CC polynucleotide and polypeptide sequences used in the exemplification
 CC of the present invention.

XX
 SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 22; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2,2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGTTPVPLLLVGVGVEKFTMTVLGIG 60
 DB 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGTTPVPLLLVGVGVEKFTMTVLGIG 60

QY 61 PVILGLVCPVLLGSASDHWRGRRPPTIWAISLIGLISLFLIPRAGWLAGLLCPDPRLL 120
 DB 61 PVILGLVCPVLLGSASDHWRGRRPPTIWAISLIGLISLFLIPRAGWLAGLLCPDPRLL 120

QY 121 ELALLILGVGLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVGLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTSAALPYLGTQBECLFGLLTILFTCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 DB 181 IDWDTSAALPYLGTQBECLFGLLTILFTCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY 241 CCPCARLAFRNLGALLPRHLCCRMPTLRLLFVAELCSWMLMTFTLFTYDFVGEGL 300
 DB 241 CCPCARLAFRNLGALLPRHLCCRMPTLRLLFVAELCSWMLMTFTLFTYDFVGEGL 300

QY 301 YQGVPRAEPTGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
 DB 301 YQGVPRAEPTGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRFQTRAVYLASVA 360

QY 361 APFVAAGATCLSHSVAVVTASAALTGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 APFVAAGATCLSHSVAVVTASAALTGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 DB 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAAGLGLVAIYFATQ 540
 DB |||||
 QY 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAAGLGLVAIYFATQ 540
 DB |||||
 QY 541 VVFDKSLAKYSA 553
 DB |||||
 QY 541 VVFDKSLAKYSA 553
 DB |||||

RESULT 9
 AAG62150
 ID AAG62150 standard; Protein; 553 AA.
 XX
 AC AAG62150;
 XX
 DT 06-JUL-2001 (first entry)
 XX
 DE Human P501S inventive antigen SEQ ID NO: 333.
 XX
 KW Human; mouse; immunotherapy; cancer; leukaemia; WTI; Wilm's tumour gene;
 KW chromosome 11p13; zinc finger transcription factor.
 XX
 OS Homo sapiens.
 XX
 FN WO200125273-A2.
 XX
 PD 12-APR-2001.
 XX
 PF 04-OCT-2000; 2000WO-US27465.
 XX
 PR 04-OCT-1999; 99US-0157459.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Xu J, Cheever MA, Reed SG;
 XX
 DR WPI; 2001-328324/34.
 XX
 XX Polypeptide comprising part of the Wilm's Tumour gene product sequence is
 PT used in the diagnosis and treatment of malignant diseases e.g. leukemia
 PT and cancer associated with WTI
 XX
 PS Disclosure; Page 212-213; 228pp; English.
 XX
 CC The present invention describes compositions comprising peptides derived
 CC from the Wilm's tumour protein WTI and methods for their use in treating
 CC malignant diseases. Peptides derived from both the murine and human WTI
 CC proteins are provided. The human WTI gene is found on chromosome 11p13,
 CC and the protein was shown to be a zinc finger transcription factor. The
 CC -immunogenic peptides of the invention are particularly useful in the
 CC diagnosis and treatment of cancer and leukaemia. The present sequence is
 CC a polypeptide described in the exemplification of the invention.
 XX
 SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 22; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRHRKAQLLLVNLTLFGLVCLAAAGITVVPPLLEVGVEEKFMTMVLGIG 60
 DB |||||
 QY 61 PVLGLVCVPLLGASDHWGRGRRRPFIALSIGILLSFLIPRAGWLAGLCPDRPL 120
 DB |||||
 QY 61 PVLGLVCVPLLGASDHWGRGRRRPFIALSIGILLSFLIPRAGWLAGLCPDRPL 120
 DB |||||
 QY 121 ELALLIGVGLLDFCGQVCFTPLALLSDLPDPDHCQAVSVVAFMTSLGCGLYLAPA 180
 DB |||||
 QY 121 ELALLIGVGLLDFCGQVCFTPLALLSDLPDPDHCQAVSVVAFMTSLGCGLYLAPA 180
 DB |||||
 QY 181 IDWTSALAPYLGTQOECLFGLLTILFILTCAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
 DB |||||

DB 181 IDWTSALAPYLGTQOECLFGLLTILFILTCAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
 QY |||||
 DB 241 CCPCARLAFAFNIGLALLPRLHQLCCRMPTLRLRLFAELCSMMALMTTLFYDFVGEGL 300
 DB |||||
 DB 241 CCPCARLAFAFNIGLALLPRLHQLCCRMPTLRLRLFAELCSMMALMTTLFYDFVGEGL 300
 QY |||||
 DB 301 YQGVPRAPSGTEARRHYDEGVRMGSIGLFLQCAISLVSFLVMDRLVQRFGTTRAVYLASVA 360
 DB |||||
 DB 301 YQGVPRAPSGTEARRHYDEGVRMGSIGLFLQCAISLVSFLVMDRLVQRFGTTRAVYLASVA 360
 QY |||||
 DB 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB |||||
 DB 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY |||||
 DB 421 ASSSDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSRVVVGEPTA 480
 DB |||||
 DB 421 ASSSDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSRVVVGEPTA 480
 QY |||||
 DB 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAAGLGLVAIYFATQ 540
 DB |||||
 QY 541 VVFDKSLAKYSA 553
 DB |||||
 DB 541 VVFDKSLAKYSA 553
 DB |||||

RESULT 10
 AAB74800
 ID AAB74800 standard; Protein; 553 AA.
 XX
 AC AAB74800;
 XX
 DT 14-JUN-2001 (first entry)
 XX
 DE Prostate tumour antigen predicted amino acid sequence for L1-12.
 XX
 KW Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
 KW prostate cancer; immunogenic; cytostatic; vaccine.
 XX
 OS Homo sapiens.
 XX
 FN WO200125272-A2.
 XX
 PD 12-APR-2001.
 XX
 PF 04-OCT-2000; 2000WO-US27464.
 XX
 PR 04-OCT-1999; 99US-0157455.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Skeiky YAW, Reed SG, Cheever MA;
 XX
 DR WPI; 2001-245062/25.
 XX
 N-PSDB; AAH02530.
 XX
 PT Prostate specific protein and its encoding polynucleotide, useful for
 PT the treatment and diagnosis of prostate cancer
 XX
 PS Claim 3; Page 157-158; 276pp; English.
 XX
 CC The present invention describes an isolated polypeptide (I) comprising
 CC at least an immunogenic portion of a prostate tumour antigen protein or
 CC its variant. (I) have cytostatic activity and can be used in vaccine
 CC production. (II), prostate tumour antigen polynucleotides, an antigen
 CC presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
 CC pharmaceutical composition containing (I) are useful for inhibiting the
 CC development of cancer in a patient. Antibodies specific for prostate
 CC specific proteins and oligonucleotides that hybridise to a
 CC polynucleotide that encodes a prostate specific protein are useful
 CC for detecting the presence or absence of a cancer or monitoring the
 CC progression the progression of a cancer, especially prostate cancer.

XX ABG76665;
 XX 05-NOV-2002 (first entry)
 XX Prostate tumour protein #3.
 XX Human; prostate tumour; immunotherapy; prostate cancer.
 XX Homo sapiens.
 XX US2002081580-A1:
 XX 27-JUN-2002.
 XX 25-FEB-1998; 98US-0030606.
 XX 25-FEB-1997; 97US-0806596.
 XX 01-AUG-1997; 97US-0904809.
 XX 09-FEB-1998; 98US-0020747.
 XX (XUJ/) XU J.
 XX (DILL/) DILLON D C.
 XX Xu J, Dillon DC;
 XX WPI; 2002-607662/65.
 XX Detecting prostate cancer comprises contacting a sample with an agent
 XX capable of binding to a polypeptide with an immunogenic portion of a
 XX prostate protein, oligonucleotide primers or a probe specific for DNA
 XX encoding the polypeptide
 XX Example 1; Page 65-67; 111pp; English.
 XX The invention relates to a method of detecting prostate cancer by
 XX contacting a biological sample from a patient with: (a) a binding agent
 XX that binds to a polypeptide having an immunogenic portion of a prostate
 XX protein or its variant; (b) 2 oligonucleotide primers, where 1 of the
 XX oligonucleotides is specific for a DNA encoding the polypeptide of (a);
 XX or (c) an oligonucleotide probe specific for a DNA molecule encoding the
 XX polypeptide of (a). The method and polypeptides are useful for
 XX diagnosing, treating, particularly by immunotherapy, monitoring the
 XX progression, and inhibiting the development of prostate cancer in a
 XX patient. The polypeptides may be used to generate antibodies useful for
 XX the diagnosis and monitoring of prostate cancer. ABG76663-ABG76669
 XX represent human prostate tumour protein sequences of the invention.
 XX Sequence 553 AA;
 XX Query Match 100.0%; Score 2861; DB 23; Length 553;
 XX Best Local Similarity 100.0%; Pred. No. 2.2e-269;
 XX Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVQRLWVGRLLRHRKAQLLNLLTFGLVCLAGITVPPLLLEVGVEEKFMTVLGIG 60
 DB 1 MVQRLWVGRLLRHRKAQLLNLLTFGLVCLAGITVPPLLLEVGVEEKFMTVLGIG 60
 QY 61 PVGLHVCVPLIGSADHWGRYGRRRPTIWLISGLILSLFLIPRAGWLAGLLCPDRPL 120
 DB 61 PVGLHVCVPLIGSADHWGRYGRRRPTIWLISGLILSLFLIPRAGWLAGLLCPDRPL 120
 QY 121 ELALLILGVLLDFCGQVCFTEPLLEALLSDLRDPDHCQAVSVYAFMISLGGCLYLPA 180
 DB 121 ELALLILGVLLDFCGQVCFTEPLLEALLSDLRDPDHCQAVSVYAFMISLGGCLYLPA 180
 QY 181 IDWDTLSALAPYLGTEECFLGILLTIFLTVAATLLVAEEAALGTEPAEGLSAPLSPH 240
 DB 181 IDWDTLSALAPYLGTEECFLGILLTIFLTVAATLLVAEEAALGTEPAEGLSAPLSPH 240
 QY 241 CCPCARLAFNRLGALLPRLHQLCCMRPTRLRLRFLVABLCSWMLMTTFLFYDFVSGGL 300
 DB 241 CCPCARLAFNRLGALLPRLHQLCCMRPTRLRLRFLVABLCSWMLMTTFLFYDFVSGGL 300

QY 301 YQGVPAEPGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
 DB 301 YQGVPAEPGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
 QY 361 AFPVAAGATCLSHSVAVVTASAAALTGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 AFPVAAGATCLSHSVAVVTASAAALTGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY 421 ASSEDSLMTSFLPGPKPGAPFNGHVAGSGGLPPPPALCGASACDVSRVVVGEPTEA 480
 DB 421 ASSEDSLMTSFLPGPKPGAPFNGHVAGSGGLPPPPALCGASACDVSRVVVGEPTEA 480
 QY 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMYSAAGLGLVAIFYATQ 540
 DB 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMYSAAGLGLVAIFYATQ 540
 QY 541 VVFDKSDLAKYSA 553
 DB 541 VVFDKSDLAKYSA 553
 RESULT 13
 ABB77575
 ID ABB77575 standard; Protein; 553 AA.
 XX AC ABB77575;
 XX DT 30-AUG-2002 (first entry)
 XX DE Human mast cell related splice variant protein MC14 SEQ ID NO 13.
 XX KW Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
 XX KW vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma.
 XX OS Homo sapiens.
 XX PN WO200246389-A2.
 XX PD 13-JUN-2002.
 XX PF 07-DEC-2001; 2001WO-US46180.
 XX PR 08-DEC-2000; 2000US-251835P.
 XX PR 14-MAR-2001; 2001US-275479P.
 XX PR 28-MAR-2001; 2001US-279115P.
 XX PR 02-APR-2001; 2001US-280143P.
 XX PA (UNIO) UCB SA.
 XX PI Nocka K, Pirozzi G, Einstein R;
 XX WPI; 2002-508560/54.
 XX DR N-PSDB; AEN81324.
 XX PT Novel isolated nucleic acids that are differentially expressed in mast
 XX cells in patients with allergic hypersensitivity, encoding proteins
 XX associated with mast cell regranulation and allergic hypersensitivity
 XX Claim 31; Page 117-119; 119pp; English.
 XX The invention relates to isolated nucleic acid (AEN81319-AEN81324),
 XX corresponding to genes differentially expressed in mast cells following
 XX activation or in patients with allergic hypersensitivity disease, (I)
 XX that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of
 XX (II) if at least 6 amino acids. (II) is useful for identifying binding
 XX partners. (I) or (II) is useful for diagnosing or treating a disease
 XX state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
 XX urticaria or atopic dermatitis or mastocytosis) in a subject which
 XX involves determining the level of expression of (I) or (II). A computer
 XX system, comprising a database containing information identifying the
 XX expression level in a tissue or at least one mast cell of (I), is useful

CC for presenting information to identify the relative expression level of
CC (I) (II) is used as a marker to detect, diagnose or identify an allergic
CC response in a patient. The protein can also serve as a target that
CC modulate gene expression or activity and as an antigen to raise
CC polyclonal or monoclonal antibodies. (II) is useful for identifying
CC agents that modulate expression of the protein or agents, such as
CC agonists or antagonists. The agonists or antagonists are useful for
CC modulating biological activity and function of (II) and thus are useful
CC for alleviating disease conditions such as allergic hypersensitivity,
CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis.
XX
SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 23; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVQRLVWSRLRRHKAQLLLVNLITFGLVCLAAAGTTPVPLLEVGVEEKFMTWLGIG 60
DB 1 MVQRLVWSRLRRHKAQLLLVNLITFGLVCLAAAGTTPVPLLEVGVEEKFMTWLGIG 60
QY 61 PVGLVVCVPLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
DB 61 PVGLVVCVPLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLITLGVLLDFCGOVCFPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLITLGVLLDFCGOVCFPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTSAALPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSLPH 240
DB 181 IDWDTSAALPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSLPH 240
QY 241 CCPCARLAFRNLGALLPRHLQCCMPRTLRLFVAELCSWMLMTFTLYTDFVGEGL 300
DB 241 CCPCARLAFRNLGALLPRHLQCCMPRTLRLFVAELCSWMLMTFTLYTDFVGEGL 300
QY 301 YQGVPAEPCTEARRHYDEGVNMGSLGLFLOCAISLVFSIAMDRLVQRECTRAVYLASVA 360
DB 301 YQGVPAEPCTEARRHYDEGVNMGSLGLFLOCAISLVFSIAMDRLVQRECTRAVYLASVA 360
QY 361 APFVAAGATCLSHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APFVAAGATCLSHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGSGSLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGSGSLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQVTAVMVSAAGILVAIYFATQ 540
DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQVTAVMVSAAGILVAIYFATQ 540
QY 541 VFDPKSDLAKYSA 553
DB 541 VFDPKSDLAKYSA 553

RESULT 14

ABG61900
ID ABG61900 standard; Protein; 553 AA.

XX ABG61900;

XX 15-AUG-2002 (first entry)

DT Prostate cancer-associated protein #101.

DE Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

XX Mammalia.

XX WO200230268-A2.

XX 18-APR-2002.
XX 12-OCT-2001; 2001WO-US32045.
XX 13-OCT-2000; 2000US-0687576.
XX 08-DEC-2000; 2000US-0733288.
XX 24-JAN-2001; 2001US-263957P.
XX 16-MAR-2001; 2001US-276791P.
XX 16-MAR-2001; 2001US-276888P.
XX 06-APR-2001; 2001US-281922P.
XX 24-APR-2001; 2001US-286214P.
XX 30-APR-2001; 2001US-0847046.
XX 04-MAY-2001; 2001US-288589P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX WPI; 2002-471335/50.
XX N-PSDB; ABK92217.
XX Detecting a prostate cancer-associated transcript in a cell in a
XX patient, useful for diagnosing prostate cancer (PC) or screening
XX modulators of PC, by determining if prostate cancer-associated genes
XX are expressed in a prostate tissue -
XX Claim 27; Page 386; 436pp; English.
XX The present invention relates to methods of detecting a prostate
XX cancer-associated transcript in a cell from a patient. The method
XX comprises contacting a biological sample from the patient with
XX prostate cancer-associated polynucleotides (designated PC genes) that
XX selectively hybridise to a sequence that is at least 80% identical
XX to them. The prostate cancer-associated polynucleotide sequences
XX are differentially expressed in prostate tumour tissue or in
XX prostate cancer and are derived from the tissues of various
XX organisms such as humans or other mammals (e.g. mice, sheep and dogs).
XX The methods of the invention are useful for diagnosing and treating
XX prostate cancer in mammals. The prostate cancer-associated genes are
XX useful for diagnosing or treating prostate cancer, as well as for
XX identifying modulators of prostate cancer or agents that inhibit
XX prostate cancer. The nucleic acid sequences are particularly useful
XX in gene therapy, as a vaccine or in antisense applications.
XX ABG61800-ABG61944 represent prostate cancer-associated proteins.

SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 23; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLVWSRLRRHKAQLLLVNLITFGLVCLAAAGTTPVPLLEVGVEEKFMTWLGIG 60

DB 1 MVQRLVWSRLRRHKAQLLLVNLITFGLVCLAAAGTTPVPLLEVGVEEKFMTWLGIG 60

QY 61 PVGLVVCVPLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

DB 61 PVGLVVCVPLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLITLGVLLDFCGOVCFPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180

DB 121 ELALLITLGVLLDFCGOVCFPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTSAALPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSLPH 240

DB 181 IDWDTSAALPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSLPH 240

QY 241 CCPCARLAFRNLGALLPRHLQCCMPRTLRLFVAELCSWMLMTFTLYTDFVGEGL 300

DB 241 CCPCARLAFRNLGALLPRHLQCCMPRTLRLFVAELCSWMLMTFTLYTDFVGEGL 300

```
Qy 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Qy 361 AFPVAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 AFPVAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEA 480
Db 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEA 480
Qy 481 RVVPGRGICLDLAIDLDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAGGLVAIYFATQ 540
Db 481 RVVPGRGICLDLAIDLDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAGGLVAIYFATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 15
ABB95222
ID ABB95222 standard; Protein; 553 AA.
XX AC
XX ABB95222;
DT 19-JUL-2002 (first entry)
XX Human L1-12 protein SEQ ID NO 113.
DE Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW gene therapy.
XX Homo sapiens.
XX US2002022248-A1.
PD 21-FEB-2002.
XX 12-JAN-2001; 2001US-0759143.
XX 25-FEB-1997; 97US-0806099.
PR 01-AUG-1997; 97US-0904804.
PR 09-FEB-1998; 98US-0020956.
PR 25-FEB-1998; 98US-0030607.
PR 14-JUL-1998; 98US-0115453.
PR 23-SEP-1998; 98US-0159812.
PR 15-JAN-1999; 99US-0232149.
PR 09-APR-1999; 99US-0288946.
PR 13-JUL-1999; 99US-0352616.
PR 12-NOV-1999; 99US-0439313.
PR 18-NOV-1999; 99US-0443686.
PR 14-JAN-2000; 2000US-0483672.
PR 27-MAR-2000; 2000US-0536857.
PR 09-MAY-2000; 2000US-0568100.
PR 12-MAY-2000; 2000US-0570737.
PR 13-JUN-2000; 2000US-0593793.
PR 27-JUN-2000; 2000US-0605783.
PR 10-AUG-2000; 2000US-0636215.
PR 29-AUG-2000; 2000US-0651236.
PR 06-SEP-2000; 2000US-0657279.
PR 02-OCT-2000; 2000US-0679426.
PR 10-OCT-2000; 2000US-0685166.
XX (XUJJ/) XU J.
XX (DILL/) DILLON D C.
PA (MITC/) MITCHAM J L.
PA (HARL/) HARLOCKER S L.
PA (JIAN/) JIANG Y.
PA (KALO/) KALOS M D.
PA (FANG/) FANGER G R.
PA (RETT/) RETTER M W.
```

```
PA (STOL/) STOLK J A.
PA (DAYC/) DAY C H.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
PA (LISX/) LI S X.
PA (WANG/) WANG A.
PA (SKEI/) SKEIKY Y A W.
PA (HEPL/) HEPLER W T.
PA (HEND/) HENDERSON R A.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX WPI; 2002-255649/30.
XX
PT New prostate-specific polynucleotides for diagnosing and treating
PT diseases, in particular prostate cancer, and as markers for the
PT progression of cancer
XX
XX Claim 2; SEQ ID NO 113; 87pp; English.
XX
CC The present invention provides prostate-specific coding sequences and
CC their encoded proteins. These can be used in the diagnosis and treatment
CC of cancers, particularly prostate cancer. The present sequence is a
CC protein described in the invention.
XX
SQ Sequence 553 AA;
Query Match 100.0%; Score 2861; DB 23; Length 553;
Best Local Similarity 100.0%; Pred. NO. 2.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVQRLWVSRLLRHRAQALLVNLTFGLVCLAAAGITVPPILLVGVVEKEKMTWLIGT 60
Db 1 MVQRLWVSRLLRHRAQALLVNLTFGLVCLAAAGITVPPILLVGVVEKEKMTWLIGT 60
Qy 61 PVLGLVCPVLIGSASDHWGRYGRRRPFTWALSGLISLFLI PRAGWLAGLLCPDPRPL 120
Db 61 PVLGLVCPVLIGSASDHWGRYGRRRPFTWALSGLISLFLI PRAGWLAGLLCPDPRPL 120
Qy 121 ELALLILGVGLDFCGQVCFTLEALLSDLPDPDHCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGQVCFTLEALLSDLPDPDHCQAYSVYAFMISLGGCLGYLLPA 180
Qy 181 IDWDTSAALPYLGTQECFLGILLTLIFLTCVAATLLVAEEAALGTEPAEGLSAPLSPH 240
Db 181 IDWDTSAALPYLGTQECFLGILLTLIFLTCVAATLLVAEEAALGTEPAEGLSAPLSPH 240
Qy 241 CCPCRARLAFRNLGALLPRLHQLCCRMPTLRLRFLVAELCSWMLMTFTLFTYDFVGEGL 300
Db 241 CCPCRARLAFRNLGALLPRLHQLCCRMPTLRLRFLVAELCSWMLMTFTLFTYDFVGEGL 300
Qy 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Qy 361 APPVAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APPVAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEA 480
Db 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEA 480
Qy 481 RVVPGRGICLDLAIDLDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAGGLVAIYFATQ 540
Db 481 RVVPGRGICLDLAIDLDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAGGLVAIYFATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553
```


Search completed: December 3, 2003, 17:19:18
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 17:18:21 ; Search time 36 Seconds
(without alignments)
2856.916 Million cell updates/sec

Title: US-09-593-793a-113

Perfect score: 2861

Sequence: 1 MVQRLWVSRLLRHRAQLL.....AIYFATQVVFVDSKLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2861	100.0	553	9	US-09-745-288-101
2	2861	100.0	553	9	US-09-838-785-2
3	2861	100.0	553	9	US-09-759-143-113
4	2861	100.0	553	9	US-09-780-669-113
5	2861	100.0	553	9	US-09-030-606-113
6	2861	100.0	553	9	US-09-822-827-113
7	2861	100.0	553	9	US-09-115-453-113
8	2861	100.0	553	10	US-09-232-880-113
9	2861	100.0	553	10	US-09-895-793-113
10	2861	100.0	553	10	US-09-895-814-113
11	2861	100.0	553	12	US-10-144-678A-113
12	2861	100.0	553	12	US-10-005-907-13
13	2861	100.0	553	12	US-10-294-025-113
14	2861	100.0	553	14	US-10-012-896-113
15	2861	100.0	553	15	US-10-010-940-113

16	2861	100.0	1079	9	US-09-822-827-947	Sequence 947, App
17	2861	100.0	1079	10	US-09-895-793-947	Sequence 947, App
18	2601	90.9	710	15	US-10-296-770-4	Sequence 4, Appli
19	1696	59.3	359	9	US-09-822-827-974	Sequence 974, App
20	1696	59.3	359	10	US-09-895-793-974	Sequence 974, App
21	1677.5	58.6	305	15	US-10-296-770-5	Sequence 5, Appli
22	1517	53.0	305	12	US-10-144-678A-1029	Sequence 1029, Ap
23	1517	53.0	305	12	US-10-294-025-1029	Sequence 1029, Ap
24	1417.5	49.5	371	9	US-09-759-143-708	Sequence 708, App
25	1417.5	49.5	371	9	US-09-780-669-708	Sequence 708, App
26	1417.5	49.5	371	9	US-09-822-827-708	Sequence 708, App
27	1417.5	49.5	371	10	US-09-895-793-708	Sequence 708, App
28	1417.5	49.5	371	10	US-09-895-814-708	Sequence 708, App
29	1417.5	49.5	371	12	US-10-144-678A-708	Sequence 708, App
30	1417.5	49.5	371	12	US-10-294-025-708	Sequence 708, App
31	1417.5	49.5	371	14	US-10-012-896-708	Sequence 708, App
32	1403.5	49.1	400	9	US-09-759-143-852	Sequence 852, App
33	1403.5	49.1	400	9	US-09-780-669-852	Sequence 852, App
34	1403.5	49.1	400	9	US-09-822-827-852	Sequence 852, App
35	1403.5	49.1	400	10	US-09-895-793-852	Sequence 852, App
36	1403.5	49.1	400	10	US-09-895-814-852	Sequence 852, App
37	1403.5	49.1	400	12	US-10-144-678A-852	Sequence 852, App
38	1403.5	49.1	400	12	US-10-294-025-852	Sequence 852, App
39	1403.5	49.1	400	14	US-10-012-896-852	Sequence 1028, Ap
40	1197	41.8	246	12	US-10-294-025-1028	Sequence 1028, Ap
41	1197	41.8	246	12	US-10-005-907-5	Sequence 5, Appli
42	1014.5	35.5	326	12	US-10-005-907-5	Sequence 5, Appli
43	947	33.1	355	12	US-10-144-678A-1011	Sequence 1011, Ap
44	947	33.1	355	12	US-10-294-025-1011	Sequence 1011, Ap
45	947	33.1	355	14	US-10-012-896-1011	Sequence 1011, Ap

ALIGNMENTS

RESULT 1

US-09-745-288-101
; Sequence 101, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446D1
; CURRENT APPLICATION NUMBER: US/09/745,288
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 101
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-745-288-101

Query Match	100.0%;	Score 2861;	DB 9;	Length 553;
Best Local Similarity	100.0%;	Pred. No. 1.9e-245;		
Matches 553;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MYQRLWVSRLLRHRAQALLNNLLTFGLEVCIAAGITYYPPLLLELVGVBEKFTMWLGIG	60	
Db	1	MYQRLWVSRLLRHRAQALLNNLLTFGLEVCIAAGITYYPPLLLELVGVBEKFTMWLGIG	60	
QY	61	PVLGLVCPVLLGSADHWRGRRRPFITWALSGLILLSLFLIPRAGWLAGLLCPDPRPL	120	
Db	61	PVLGLVCPVLLGSADHWRGRRRPFITWALSGLILLSLFLIPRAGWLAGLLCPDPRPL	120	
QY	121	ELALLILGVLLDFGQVCFTPLEALLSLFRDPDHCQAYSYYAFMISLGGCGLYLLPA	180	
Db	121	ELALLILGVLLDFGQVCFTPLEALLSLFRDPDHCQAYSYYAFMISLGGCGLYLLPA	180	
QY	181	IDWDTSALAPYLGTOBECLFGLLTITFLTCVAATLIVAEAAALGPTPEAGLSAPLSLPH	240	

```
Db 181 IDWDTAALPYLGTQBECLFGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Qy 241 CCPCARLAFRNLGALLPRHLQCCMRPRTLRLFLVAELCSMMALMTFTFLFYTDFVGEGL 300
Db 241 CCPCARLAFRNLGALLPRHLQCCMRPRTLRLFLVAELCSMMALMTFTFLFYTDFVGEGL 300
Qy 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
Qy 361 AFPVAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 AFPVAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Qy 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAYIFATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAYIFATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553
```

RESULT 2

US-09-838-785-2

```
; Sequence 2, Application US/09838785
; Patent No. US20020009455A1
; GENERAL INFORMATION:
; APPLICANT: Lau, Ted
; APPLICANT: Lin, Rick
; APPLICANT: Parkes, Debbie
; APPLICANT: Parry, Gordon
; APPLICANT: Schneider, Douglas
; APPLICANT: Steinbrecher, Renate
; APPLICANT: Van Heut, Pam T
; APPLICANT: Wu, John
; TITLE OF INVENTION: DNA Encoding a No. US20020009455A1el PROST 03
; FILE REFERENCE: 51831AUSM1
; CURRENT APPLICATION NUMBER: US/09/838,785
; CURRENT FILING DATE: 2001-04-20
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-838-785-2
```

```
Query Match 100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVORLWVSRLRRHKAQALLVNLTFGLVCLAAAGTTPVPLLEVGVBKEKFTMWLGIG 60
Db 1 MVORLWVSRLRRHKAQALLVNLTFGLVCLAAAGTTPVPLLEVGVBKEKFTMWLGIG 60
Qy 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
Db 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
Qy 121 ELALLIIGVLLDFCGQVCFTPEALLSDLFROPDHCRCQAYSVYAFMISGCGLYLLPA 180
Db 121 ELALLIIGVLLDFCGQVCFTPEALLSDLFROPDHCRCQAYSVYAFMISGCGLYLLPA 180
Qy 181 IDWDTAALPYLGTQBECLFGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
```

```
Db 181 IDWDTAALPYLGTQBECLFGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Qy 241 CCPCARLAFRNLGALLPRHLQCCMRPRTLRLFLVAELCSMMALMTFTFLFYTDFVGEGL 300
Db 241 CCPCARLAFRNLGALLPRHLQCCMRPRTLRLFLVAELCSMMALMTFTFLFYTDFVGEGL 300
Qy 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
Qy 361 AFPVAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 AFPVAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Qy 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAYIFATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAYIFATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553
```

RESULT 3

```
US-09-759-143-113
; Sequence 113, Application US/09759143
; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-759-143-113
```

```
Query Match 100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVORLWVSRLRRHKAQALLVNLTFGLVCLAAAGTTPVPLLEVGVBKEKFTMWLGIG 60
Db 1 MVORLWVSRLRRHKAQALLVNLTFGLVCLAAAGTTPVPLLEVGVBKEKFTMWLGIG 60
Qy 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
Db 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
```

121 ELALLILVGLLDFCCQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 121 ELALLILVGLLDFCCQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 181 IDWDTGALAPYLGTOBECLFGLLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 181 IDWDTGALAPYLGTOBECLFGLLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 241 CCPCARLAFRNLGALLPRLHOLCCMRPRTLRLFLVAELCSWMLMTFTFLFYDFVGEGL 300
 241 CCPCARLAFRNLGALLPRLHOLCCMRPRTLRLFLVAELCSWMLMTFTFLFYDFVGEGL 300
 301 YQGVPAEPTGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 301 YQGVPAEPTGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 361 APVAAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 361 APVAAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 481 RVVPGRGICLDLAILDSAPFLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAYIFATQ 540
 481 RVVPGRGICLDLAILDSAPFLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAYIFATQ 540
 541 VVFDKSDLAKYSA 553
 541 VVFDKSDLAKYSA 553

RESULT 4

US-09-780-669-113
 ; Sequence 113, Application US/09780669
 ; Patent No. US20020051977A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yuqui
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedrick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Li, Samuel
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Hepler, William
 ; APPLICANT: Hural, John
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Houghton, Raymond L.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCE: 210121.427C24
 ; CURRENT APPLICATION NUMBER: US/09/780,669
 ; CURRENT FILING DATE: 2001-02-09
 ; NUMBER OF SEQ ID NOS: 943
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 113
 ; LENGTH: 553
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; US-09-780-669-113
 Query Match 100.0%; Score 2861; DB 9; Length 553;

Best Local Similarity 100.0%; Pred. No. 1.9e-245;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MYORLWVSELLRRKRAQALLVNLTLFGLVCLAAAGITVYVPPLLLVGVBEKFWMTVLGIG 60
 DB 1 MYORLWVSELLRRKRAQALLVNLTLFGLVCLAAAGITVYVPPLLLVGVBEKFWMTVLGIG 60
 QY 61 PVLGLVCVPLLGASADHMRGRRPFTIWAISLGLTSLFLIPRAGWLAGLCPDPRPL 120
 DB 61 PVLGLVCVPLLGASADHMRGRRPFTIWAISLGLTSLFLIPRAGWLAGLCPDPRPL 120
 QY 121 ELALLILVGLLDFCCQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILVGLLDFCCQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDWDTGALAPYLGTOBECLFGLLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 DB 181 IDWDTGALAPYLGTOBECLFGLLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 QY 241 CCPCARLAFRNLGALLPRLHOLCCMRPRTLRLFLVAELCSWMLMTFTFLFYDFVGEGL 300
 DB 241 CCPCARLAFRNLGALLPRLHOLCCMRPRTLRLFLVAELCSWMLMTFTFLFYDFVGEGL 300
 QY 301 YQGVPAEPTGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 DB 301 YQGVPAEPTGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 QY 361 APVAAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 APVAAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 QY 481 RVVPGRGICLDLAILDSAPFLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAYIFATQ 540
 DB 481 RVVPGRGICLDLAILDSAPFLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAYIFATQ 540
 QY 541 VVFDKSDLAKYSA 553
 DB 541 VVFDKSDLAKYSA 553

RESULT 5

US-09-030-606-113
 ; Sequence 113, Application US/09030606
 ; Patent No. US20020081580A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF PROSTATE CANCER AND METHODS FC
 ; NUMBER OF SEQUENCES: 224
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEED and BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/030,606
 ; FILING DATE: 25-FEB-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Maki, David J.
 ; REGISTRATION NUMBER: 31,392
 ; REFERENCE/DOCKET NUMBER: 210121.428C3

TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 553 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-09-030-606-113

Query Match 100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLEVCLAAAGITYVPPLLLLEVGVEKFMVWLIG 60
DB 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLEVCLAAAGITYVPPLLLLEVGVEKFMVWLIG 60

QY 61 PVGLVCVPLIGSASDHWRGRRRPFIIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
DB 61 PVGLVCVPLIGSASDHWRGRRRPFIIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120

QY 121 ELALITLGVGLDFCGQVCFPLEALLSDLFDRPDHCRQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALITLGVGLDFCGQVCFPLEALLSDLFDRPDHCRQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWTSALAPYLGTOBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
DB 181 IDWTSALAPYLGTOBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY 241 CCPCARLAFRNIGALLPRHLQCCMPRTLRFLVAELCSWMLMTFTLFYDFVGEGL 300
DB 241 CCPCARLAFRNIGALLPRHLQCCMPRTLRFLVAELCSWMLMTFTLFYDFVGEGL 300

QY 301 YQGVPAEPTGTEARRHYDEGVRMSGLFLQCAISLVFSLVMDRLVQRTGTRAVYLA 360
DB 301 YQGVPAEPTGTEARRHYDEGVRMSGLFLQCAISLVFSLVMDRLVQRTGTRAVYLA 360

QY 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTG 420
DB 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTG 420

QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPT 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPT 480

QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAAGLGLVAYIFATQ 540
DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAAGLGLVAYIFATQ 540

QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 6

US-09-822-827-113
Sequence 113, Application US/09822827
Patent No. US20020081680A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.534C1
CURRENT APPLICATION NUMBER: US/09/822,827
CURRENT FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 982
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 113

LENGTH: 553
TYPE: PRT
ORGANISM: Homo sapien
US-09-822-827-113

Query Match 100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLEVCLAAAGITYVPPLLLLEVGVEKFMVWLIG 60
DB 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLEVCLAAAGITYVPPLLLLEVGVEKFMVWLIG 60

QY 61 PVGLVCVPLIGSASDHWRGRRRPFIIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
DB 61 PVGLVCVPLIGSASDHWRGRRRPFIIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120

QY 121 ELALITLGVGLDFCGQVCFPLEALLSDLFDRPDHCRQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALITLGVGLDFCGQVCFPLEALLSDLFDRPDHCRQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWTSALAPYLGTOBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
DB 181 IDWTSALAPYLGTOBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY 241 CCPCARLAFRNIGALLPRHLQCCMPRTLRFLVAELCSWMLMTFTLFYDFVGEGL 300
DB 241 CCPCARLAFRNIGALLPRHLQCCMPRTLRFLVAELCSWMLMTFTLFYDFVGEGL 300

QY 301 YQGVPAEPTGTEARRHYDEGVRMSGLFLQCAISLVFSLVMDRLVQRTGTRAVYLA 360
DB 301 YQGVPAEPTGTEARRHYDEGVRMSGLFLQCAISLVFSLVMDRLVQRTGTRAVYLA 360

QY 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTG 420
DB 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTG 420

QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPT 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPT 480

QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAAGLGLVAYIFATQ 540
DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAAGLGLVAYIFATQ 540

QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 7

US-09-115-453-113
Sequence 113, Application US/09115453B
Patent No. US20020090372A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND
METHODS FOR THEIR USE
FILE REFERENCE: 210121.427C4
CURRENT APPLICATION NUMBER: US/09/115,453B
CURRENT FILING DATE: 1998-07-14
NUMBER OF SEQ ID NOS: 228
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 113
LENGTH: 553
TYPE: PRT
ORGANISM: Homo sapien
US-09-115-453-113

Query Match 100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLVLLTGLFVCLAAAGITYVPPILLLEVGVEEKFTMWLGIG 60
DB 1 MVQRLWVSRLLRHRKAQLLVLLTGLFVCLAAAGITYVPPILLLEVGVEEKFTMWLGIG 60
QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLDFCGOVCFPLEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGOVCFPLEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTLSALAPYLGTQBECLFGLLTLLFLTCTVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
DB 181 IDWDTLSALAPYLGTQBECLFGLLTLLFLTCTVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
QY 241 CPCRARLAFRNLGALLPRLHOLCCMRPTRLRLFVAELCSMMALMTFTLFTYDFVGEGL 300
DB 241 CPCRARLAFRNLGALLPRLHOLCCMRPTRLRLFVAELCSMMALMTFTLFTYDFVGEGL 300
QY 301 YQGVPRAEPCTEARRHYDEGRVMSLGLFLQCAISLVSFLVMDRLVQRFCTRAVYLASVA 360
DB 301 YQGVPRAEPCTEARRHYDEGRVMSLGLFLQCAISLVSFLVMDRLVQRFCTRAVYLASVA 360
QY 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAPLLSOVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
DB 481 RVVPGRGICLDLAILDSAPLLSOVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 8

US-09-232-880-113
; Sequence 113, Application US/09232880
; Publication No. US20020182596A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF
; TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.428C6
; CURRENT APPLICATION NUMBER: US/09/232,880
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-232-880-113

Query Match 100.0%; Score 2861; DB 10; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLVLLTGLFVCLAAAGITYVPPILLLEVGVEEKFTMWLGIG 60
DB 1 MVQRLWVSRLLRHRKAQLLVLLTGLFVCLAAAGITYVPPILLLEVGVEEKFTMWLGIG 60
QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLDFCGOVCFPLEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGOVCFPLEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTLSALAPYLGTQBECLFGLLTLLFLTCTVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
DB 181 IDWDTLSALAPYLGTQBECLFGLLTLLFLTCTVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
QY 241 CPCRARLAFRNLGALLPRLHOLCCMRPTRLRLFVAELCSMMALMTFTLFTYDFVGEGL 300
DB 241 CPCRARLAFRNLGALLPRLHOLCCMRPTRLRLFVAELCSMMALMTFTLFTYDFVGEGL 300
QY 301 YQGVPRAEPCTEARRHYDEGRVMSLGLFLQCAISLVSFLVMDRLVQRFCTRAVYLASVA 360
DB 301 YQGVPRAEPCTEARRHYDEGRVMSLGLFLQCAISLVSFLVMDRLVQRFCTRAVYLASVA 360
QY 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAPLLSOVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
DB 481 RVVPGRGICLDLAILDSAPLLSOVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 9

US-09-895-793-113
; Sequence 113, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Xuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darriack
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Baesols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien

US-09-895-793-113

```

Query Match      100.0%; Score 2861; DB 10; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLVWSRLRHRKAQLLVNLTLFTGLEVCLAAGITYVPPLLLEVGVEEKEKMTWVLGIG 60
Db 1 MVQRLVWSRLRHRKAQLLVNLTLFTGLEVCLAAGITYVPPLLLEVGVEEKEKMTWVLGIG 60

QY 61 PVGLVVCVPLLGASDHWRGRRPFIWALSIGLILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVVCVPLLGASDHWRGRRPFIWALSIGLILLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPCRCQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTLSALAPYLGTQEBCLFGLLTFLITCVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240
Db 181 IDWDTLSALAPYLGTQEBCLFGLLTFLITCVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240

QY 241 CCPCARLAFRNLGALLPRLHQLCCRMPTLRLFLVAELCSNMALMTFTLFTYDFVGEGL 300
Db 241 CCPCARLAFRNLGALLPRLHQLCCRMPTLRLFLVAELCSNMALMTFTLFTYDFVGEGL 300

QY 301 YQGVPRAPGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQREGTRAVYLVA 360
Db 301 YQGVPRAPGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQREGTRAVYLVA 360

QY 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGSGGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGSGGLLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIVOLQSQTAYMVSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIVOLQSQTAYMVSAAGLGLVAIFYATQ 540

QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

```

RESULT 10

```

US-09-895-814-113
; Sequence 113, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa

```

```

; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-895-814-113

```

```

Query Match      100.0%; Score 2861; DB 10; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLVWSRLRHRKAQLLVNLTLFTGLEVCLAAGITYVPPLLLEVGVEEKEKMTWVLGIG 60
Db 1 MVQRLVWSRLRHRKAQLLVNLTLFTGLEVCLAAGITYVPPLLLEVGVEEKEKMTWVLGIG 60

QY 61 PVGLVVCVPLLGASDHWRGRRPFIWALSIGLILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVVCVPLLGASDHWRGRRPFIWALSIGLILLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPCRCQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTLSALAPYLGTQEBCLFGLLTFLITCVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240
Db 181 IDWDTLSALAPYLGTQEBCLFGLLTFLITCVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240

QY 241 CCPCARLAFRNLGALLPRLHQLCCRMPTLRLFLVAELCSNMALMTFTLFTYDFVGEGL 300
Db 241 CCPCARLAFRNLGALLPRLHQLCCRMPTLRLFLVAELCSNMALMTFTLFTYDFVGEGL 300

QY 301 YQGVPRAPGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQREGTRAVYLVA 360
Db 301 YQGVPRAPGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQREGTRAVYLVA 360

QY 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGSGGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGSGGLLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIVOLQSQTAYMVSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIVOLQSQTAYMVSAAGLGLVAIFYATQ 540

QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

```

RESULT 11

```

US-10-144-678A-113
; Sequence 113, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.

```


APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Hepler, William T.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals y de Bassols, Carlota
APPLICANT: Foy, Teresa M.
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Deng, Ta
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C28
CURRENT APPLICATION NUMBER: US/10/144, 678A
CURRENT FILING DATE: 2002-08-12
NUMBER OF SEQ ID NOS: 1033
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 113
LENGTH: 553
TYPE: PRT
ORGANISM: Homo sapiens
US-10-144-678A-113

Query Match 100.0%; Score 2861; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVORLWVSRLRRHKAQALLVNLITFGLVCLAAAGITVPPLLLLEVGVEEKFTMWLGIG 60
Db 1 MVORLWVSRLRRHKAQALLVNLITFGLVCLAAAGITVPPLLLLEVGVEEKFTMWLGIG 60

Qy 61 PVGLVCVPLLGASDHWGRGRRRPFIFWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVCVPLLGASDHWGRGRRRPFIFWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

Qy 121 ELALLILGVLLDFCGQVCFPLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVLLDFCGQVCFPLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

Qy 181 IDWTSALAPYLGTQEBCLFGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQEBCLFGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

Qy 241 CCPCRARLAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSMWALMTFTLYTDFVGEGL 300
Db 241 CCPCRARLAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSMWALMTFTLYTDFVGEGL 300

Qy 301 YQGVPAEPCTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
Db 301 YQGVPAEPCTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFQTRAVYLASVA 360

Qy 361 APVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

Qy 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

Qy 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540

Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 13
US-10-294-025-113
; Sequence 113, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun

RESULT 12
US-10-005-907-13
; Sequence 13, Application US/10005907
; Publication No. US20030166881A1
; GENERAL INFORMATION:
; APPLICANT: Union Chimique Belge, S.A.
; APPLICANT: No. US20030166881A1ka, Karl
; APPLICANT: Pirozzi, Gregory
; APPLICANT: Einstein, Richard
; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAST CELL
; TITLE OF INVENTION: ACTIVATION
; FILE REFERENCE: 053529-5005
; CURRENT APPLICATION NUMBER: US/10/005,907
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-005-907-13

Query Match 100.0%; Score 2861; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVORLWVSRLRRHKAQALLVNLITFGLVCLAAAGITVPPLLLLEVGVEEKFTMWLGIG 60
Db 1 MVORLWVSRLRRHKAQALLVNLITFGLVCLAAAGITVPPLLLLEVGVEEKFTMWLGIG 60

Qy 61 PVGLVCVPLLGASDHWGRGRRRPFIFWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVCVPLLGASDHWGRGRRRPFIFWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

Qy 121 ELALLILGVLLDFCGQVCFPLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVLLDFCGQVCFPLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

Qy 181 IDWTSALAPYLGTQEBCLFGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQEBCLFGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

Qy 241 CCPCRARLAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSMWALMTFTLYTDFVGEGL 300
Db 241 CCPCRARLAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSMWALMTFTLYTDFVGEGL 300

Qy 301 YQGVPAEPCTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
Db 301 YQGVPAEPCTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFQTRAVYLASVA 360

Qy 361 APVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

Qy 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

Qy 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540

Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

APPLICANT: Stolk, John A.
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C29
CURRENT APPLICATION NUMBER: US/10/294,025
CURRENT FILING DATE: 2002-11-12
NUMBER OF SEQ ID NOS: 1038
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 113
LENGTH: 553
TYPE: PRT
ORGANISM: Homo sapiens
US-10-294-025-113

Query Match 100.0%; Score 2861; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLLVNLLTFGLVCLAAAGITYVPPLLELVGVVEKFMVWLIG 60
DB 1 MVQRLWVSRLLRHRKAQLLLVNLLTFGLVCLAAAGITYVPPLLELVGVVEKFMVWLIG 60
QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
QY 121 ELALLILGVGLDFCGQVCFTEPLEALLSDFRDPDHCROQYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFTEPLEALLSDFRDPDHCROQYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTSLAPYLGTOBECLFGLLTFLITCVAATLLVAEEAALGPTPEAEGLSAPLSPH 240
DB 181 IDWDTSLAPYLGTOBECLFGLLTFLITCVAATLLVAEEAALGPTPEAEGLSAPLSPH 240
QY 241 CCPCARLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLFYDFVGEGL 300
DB 241 CCPCARLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLFYDFVGEGL 300
QY 301 YQGVPRAEPTGEARRHYDEGRVMSGLGLFLOCAISLVFSLVMDRLVQRFTRAVYLA 360
DB 301 YQGVPRAEPTGEARRHYDEGRVMSGLGLFLOCAISLVFSLVMDRLVQRFTRAVYLA 360
QY 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVLPKYRGDTGG 420
DB 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPT 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPT 480
QY 481 RVVPGRGICLDLAIDLDAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIYFATQ 540
DB 481 RVVPGRGICLDLAIDLDAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIYFATQ 540
QY 541 VFPDKSLAKYSA 553
DB 541 VFPDKSLAKYSA 553

RESULT 14

US-10-012-896-113
Sequence 113, Application US/10012896
Publication No. US20020183251A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillion, Devin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.

APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Meagher, Madeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C27
CURRENT APPLICATION NUMBER: US/10/012,896
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 1011
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 113
LENGTH: 553
TYPE: PRT
ORGANISM: Homo sapiens
US-10-012-896-113

Query Match 100.0%; Score 2861; DB 14; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLLVNLLTFGLVCLAAAGITYVPPLLELVGVVEKFMVWLIG 60
DB 1 MVQRLWVSRLLRHRKAQLLLVNLLTFGLVCLAAAGITYVPPLLELVGVVEKFMVWLIG 60
QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
QY 121 ELALLILGVGLDFCGQVCFTEPLEALLSDFRDPDHCROQYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFTEPLEALLSDFRDPDHCROQYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTSLAPYLGTOBECLFGLLTFLITCVAATLLVAEEAALGPTPEAEGLSAPLSPH 240
DB 181 IDWDTSLAPYLGTOBECLFGLLTFLITCVAATLLVAEEAALGPTPEAEGLSAPLSPH 240
QY 241 CCPCARLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLFYDFVGEGL 300
DB 241 CCPCARLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLFYDFVGEGL 300
QY 301 YQGVPRAEPTGEARRHYDEGRVMSGLGLFLOCAISLVFSLVMDRLVQRFTRAVYLA 360
DB 301 YQGVPRAEPTGEARRHYDEGRVMSGLGLFLOCAISLVFSLVMDRLVQRFTRAVYLA 360
QY 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVLPKYRGDTGG 420
DB 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPT 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPT 480
QY 481 RVVPGRGICLDLAIDLDAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIYFATQ 540
DB 481 RVVPGRGICLDLAIDLDAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIYFATQ 540
QY 541 VFPDKSLAKYSA 553
DB 541 VFPDKSLAKYSA 553

Search completed: December 3, 2003, 17:22:03
Job time : 37 secs

RESULT 15

US-10-010-940-113
; Sequence 113. Application US/10010940
; Publication No. US20030089062A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427D3
; CURRENT APPLICATION NUMBER: US/10/010.940
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-010-940-113

Query Match 100.0%; Score 2861; DB 15; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.9e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MVQRLWVSRLRRHKAQLLLNLTFTGLEVCCLAAGITVVPPLLEVGVEBKFTMTVLGIG	60
Db	1	MVQRLWVSRLRRHKAQLLLNLTFTGLEVCCLAAGITVVPPLLEVGVEBKFTMTVLGIG	60
Qy	61	PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCCPDRPL	120
Db	61	PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCCPDRPL	120
Qy	121	ELALLILGVGLLDFCGQVCFTEALSLDLFRPDHCRQAYSVYAFMISLGGCLGYLLPA	180
Db	121	ELALLILGVGLLDFCGQVCFTEALSLDLFRPDHCRQAYSVYAFMISLGGCLGYLLPA	180
Qy	181	IDWDTALAPYLGTOECLFGLTLFLITCVAATLLVAEEAALGPTPEAEGLSAPLSPH	240
Db	181	IDWDTALAPYLGTOECLFGLTLFLITCVAATLLVAEEAALGPTPEAEGLSAPLSPH	240
Qy	241	CCPCRARLAPRNLGALLPRHLQCCRPRTLRLFVAELCSWMAALMTFTLYTDFVGEGL	300
Db	241	CCPCRARLAPRNLGALLPRHLQCCRPRTLRLFVAELCSWMAALMTFTLYTDFVGEGL	300
Qy	301	YQGVPRAEPTGEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA	360
Db	301	YQGVPRAEPTGEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA	360
Qy	361	APVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG	420
Db	361	APVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG	420
Qy	421	ASSEDLSMTSFLCPKPGAPFNHGVGAGSGLLPPPPALCGASACDVSVRVVVGEPTEA	480
Db	421	ASSEDLSMTSFLCPKPGAPFNHGVGAGSGLLPPPPALCGASACDVSVRVVVGEPTEA	480
Qy	481	RVVPGRGICLDLAILDSAFLLSOVAPSLFMGSIQVLSQSVTAYVWSAAGLGLVAIYFATQ	540
Db	481	RVVPGRGICLDLAILDSAFLLSOVAPSLFMGSIQVLSQSVTAYVWSAAGLGLVAIYFATQ	540
Qy	541	VVFDKSDLAKYSA	553
Db	541	VVFDKSDLAKYSA	553

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 17:17:41 ; Search time 21 Seconds
(without alignments)
2532.442 Million cell updates/sec

Title: US-09-593-793a-113
Perfect score: 2861
Sequence: 1 MVQRLWVSRLRHKQAQLL.....AIYFATQVDFKSLAKYSA 553
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76.*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	347.5	12.1	515	2 T14340	sucrose-proton tra
2	337.5	11.8	523	2 T12198	sucrose transport
3	330	11.5	516	2 JQ2389	sucrose transport
4	328.5	11.5	594	2 G84441	probable sucrose-p
5	324.5	11.3	525	2 S28052	sucrose transport
6	323.5	11.3	512	2 S38196	sucrose transport
7	323.5	11.3	512	2 G86360	probable sucrose-p
8	317.5	11.1	507	2 S48789	sucrose transport
9	311	10.9	533	2 S43142	sucrose transport
10	305.5	10.7	501	2 T14339	sucrose-proton tra
11	297.5	10.4	428	2 S48788	sucrose transport
12	294	10.3	510	2 S38657	sucrose transport
13	294	10.3	523	2 S51114	sucrose-proton sym
14	292.5	10.2	512	2 P96741	probable sucrose t
15	292	10.2	537	2 T02982	probable sucrose t
16	286	10.0	474	2 A86234	hypothetical prote
17	281.5	9.8	491	2 E96691	probable sucrose-p
18	279.5	9.8	513	2 S38197	sucrose transport
19	276.5	9.7	492	2 A84520	probable sucrose-p
20	273	9.5	503	2 S52377	sucrose transport
21	235.5	8.2	553	2 T38541	probable sucrose c
22	191	6.7	452	2 F75217	hypothetical prote
23	187.5	6.6	541	2 B87532	transporter, proba
24	142	5.0	544	2 S75696	melibiose carrier
25	139.5	4.9	454	2 A75444	hypothetical prote
26	138.5	4.8	430	2 E75217	transporter PAB217
27	136	4.8	418	2 B87536	membrane protein,
28	133	4.6	389	2 G83413	probable MFS trans
29	128.5	4.5	422	2 G83503	probable MFS trans

30	128.5	4.5	594	2 A83096	probable permease
31	127.5	4.5	472	2 F82639	resistance protein
32	127	4.4	451	2 E81781	probable integral
33	126	4.4	394	2 AB3073	tetracycline resis
34	126	4.4	394	2 G98213	sugar efflux protei
35	126	4.4	451	2 A81206	drug efflux transporter
36	125.5	4.4	399	2 JQ1479	tetracycline resis
37	125	4.4	391	2 S74688	hypothetical prote
38	125	4.4	407	2 H75515	tetracycline-efflu
39	124.5	4.4	399	1 S38556	teta protein - pse
40	123.5	4.3	400	2 AG2866	MFS permease [drug
41	123.5	4.3	400	2 D97643	probable efflux pr
42	123	4.3	440	2 AD0986	hypothetical metab
43	121.5	4.2	398	2 C90349	multidrug-efflux t
44	121.5	4.2	503	2 S43017	puromycin resistan
45	121	4.2	483	2 AF2363	hypothetical prote

ALIGNMENTS

RESULT 1

T14340
sucrose-proton transport protein - carrot
N;Alternate names: sucrose/H+ symporter protein
C;Species: Daucus carota (carrot)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T14340

R;Shakya, R.; Sturm, A.
Plant Physiol. 118, 1473-1480, 1998
A;Title: Characterization of source- and sink-specific sucrose/H+ symporters from carrot
A;Reference number: Z17991; MUID:99063785; PMID:9847123
A;Accession: T14340
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-515 <SHA>
A;Cross-references: EMBL:Y16768; NID:G2969883; PIDN:CAA76369.1; PID:G2969884
A;Experimental source: cultivar Namtaise; root
C;Genetics:
A;Note: SUT2
C;Superfamily: common tobacco sucrose transport protein

Query Match 12.1%; Score 347.5; DB 2; Length 515;
Best Local Similarity 25.9%; Pred. No. 1e-18;
Matches 131; Conservative 80; Mismatches 215; Indels 79; Gaps 17;

QY	17	QLLLNLLPFGLEVLCAAGITYVPPLLLEVGVEERFMTMVLGIPVLGLVCPVLLGSASD	76
DB	34	KLVLAAIAAGVQFGWALQSLTTPVQLLGIPIHKWAAVIMLCGPISGMLVQPIVGYS	93
QY	77	HWGRYGRPRPTWALSIGLILSLFLIPRAGLAGL-----LCPDPRPELALLILGV	129
DB	94	HCQSSFGRRRPTFIAGAGCAVLSVILI---GPAADISYKAGDDMSKTLKPRATVTVFVIG	150
QY	130	GLLDPCGVCFTPLLEALLDLFR-DPDHCRQAYSVVAFMISLGGCLGY-----LL	178
DB	151	WILDVANMLQPCRCALLADLCSGDTRMRSANAFYSPFMAVGNILGVAAGSYNNLYKLF	210
QY	179	PAIDWDTSAIPYLGTQBECLFGLTLFLTCVAATLVAEEAALGPTPEAGLSAPSLS	238
DB	211	PFS--KTHACDLYCANLKSCFFIISIALIIITVVALSVVRENS--GPPDDAAGEP---	263
QY	239	PHCCPCRALAFRNIGALLPRHLQCCRNPRTLRLFVAELCSWALMTFTLYTDFVGE	298
DB	264	----PSSGKIPV--FGELIGALKOL----PRMLLLLVITCLNIAWAFPIFLDFTDMGR	313
QY	299	GLYQGVPAEPTTEARRHRYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRCFTRAVILAS	358
DB	314	EYGGT--AGQG-----KLYDQVRAGALLNSVVLGLTSIAVEYLVKRGVGVKI-LWG	366
QY	359	VAAFPVAAGATCL-----SHSV-----AVVTASA-----ALTGTFFSALQIL	395
DB	367	FYNFILAIGLVMTVVVSKVAQHREHSANGQLLPPSAGVKAGALSLSILGILPLTYSI	426

Qy 396 PYTLASLYHREKQVFLPKYRGDTGGASSEDLSMTSLFGPKPG-----APFPNGHVGA 448
Db 427 PFALASLYSSGSGAGQGLSLGVNLALIVPQMIVSVLAGPFDLSLFGGGLNLPAPVWGAISA 486
Qy 449 GSGSLLP-----PPALCGASACDVS 469
Db 487 AISGVLAIVLLPFSKDAASKLSLS 511

RESULT 2
T12198
sucrose transport protein - fava bean
C;Species: Vicia faba (fava bean)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 20-Jun-2000
C;Accession: T12198
R;Weber, H.; Borisjuk, L.; Heim, U.; Sauer, N.; Wobus, U.
Plant Cell 9, 895-908, 1997
A;Title: A role for sugar transporters during seed development: molecular characterization
A;Reference number: Z17451; MUID:97355984; PMID:9212465
A;Accession: T12198
A;Status: preliminary; translated from GS/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-523 <WEB>
A;Cross-references: EMBL:Z93774
C;Genetics:
A;Gene: suc
C;Superfamily: common tobacco sucrose transport protein

Query Match 11.8%; Score 337.5; DB 2; Length 523;
Best Local Similarity 25.8%; Pred. No. 5.9e-18;
Matches 134; Conservative 76; Mismatches 185; Indels 125; Gaps 20;

Qy 17 QLLVNLTLTFGLVCLAAITVYPPLLLEVGVBEKFTWVLGIPVLGLVCPVLLGSASD 76
Db 36 KIMVVASIAAGVQFGWALQSLTTPYVQLLGIHHTAAIYIWLCPISGMLVQVIVGHS 95
Qy 77 HWRGRRRRPFIWALSGLTLLSLFLPRAGWLAGL-----LCPDPRLELALLIGV 129
Db 96 RCTSRFRRRPFIWALSGLTLLSLFLPRAGWLAGL-----GYAADLGHSGFSDQKVRPRAGIFVWGF 152
Qy 130 GLDPCQVCFPLEALLSDFR--DPDHCQAYSVYAFMISLGGCLGYLLPAID----- 182
Db 153 WILDVANNMLOGPCRALLDGACGNQKTRNANAFSFFVAVGNVLYAGAYSKLYHVF 212
Qy 183 --WDTSAPLYGTQECLFGLTLTFLTCVAATLV--AEAAALGPTPA-----EGLSAP 235
Db 213 PFTKTRACNVYCANLKSFP--LSIALTLVATLSALIYVKETALTPEKTVVTTEDGSSG 270
Qy 236 SLSPHCCPCARL--APRNIGALLPRHLQCCRMPTLRRLFVABLCSWMALMTFLTYT 293
Db 271 GM-----PCFQSLGAPKEL-----KRPMTLLLVTCINWIAWPFLLFD 311
Qy 294 DFVGEGLYQGVPRAEPTGTEARRH--YDEGVRMGSGLFLQCAISLVFSLVMDRLVORFGR 352
Db 312 DMGKEVI-----GTVGEHAYDMGVREGALGLMNSVLGATSLGVDIILARGVG-G 363
Qy 353 AVYLASVAAPFVAAGATCLSHSVAV-----TASAALTGF 387
Db 364 VKRLGWVNFLL--AICLGLTVLTVTKLAQHSRQYAPGTGALGDLPPPSGIRKAGALTIF 420
Qy 388 TFSALQI-----LPYTLASLYHREKQVFLPKYRGDTGGASSEDLSL-----MTS 430
Db 421 SVLGVPLAITYSIPFALASIF-----SSTSGAGQGLSLGVNLALIVPQMFVS 468
Qy 431 FLPGKPG-----APFPNGHVGAAGSG-----LPPPPP 458
Db 469 VLSGPDALPGGNLPAPVGVAVNALASGILSLILPSP 508

RESULT 3

JQ2389

sucrose transport protein - potato

C;Species: Solanum tuberosum (potato)
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 21-Jul-2000
C;Accession: JQ2389; S40310
R;Riesmeier, J.W.; Hirner, B.; Frommer, W.B.
Plant Cell 5, 1591-1598, 1993
A;Title: Potato sucrose transporter expression in minor veins indicates a role in phloem
A;Reference number: JQ2389; MUID:94146554; PMID:8312741
A;Accession: JQ2389
A;Molecule type: mRNA
A;Residues: 1-516 <RIE>
A;Cross-references: EMBL:X69165; NID:G439293; PIDN:CAA48915.1; PID:G439294
A;Experimental source: cv. Desiree
A;Comment: The gene encoding for this protein is highly expressed in mature leaves.
C;Superfamily: common tobacco sucrose transport protein
C;Keywords: glycoprotein; transmembrane protein
F;31-53/Domain: transmembrane #status predicted <TM1>
F;103-122/Domain: transmembrane #status predicted <TM2>
F;141-160/Domain: transmembrane #status predicted <TM3>
F;180-200/Domain: transmembrane #status predicted <TM4>
F;226-248/Domain: transmembrane #status predicted <TM5>
F;285-304/Domain: transmembrane #status predicted <TM6>
F;331-349/Domain: transmembrane #status predicted <TM7>
F;366-385/Domain: transmembrane #status predicted <TM8>
F;409-427/Domain: transmembrane #status predicted <TM9>
F;429-448/Domain: transmembrane #status predicted <TM10>
F;429-448/Domain: transmembrane #status predicted <TM11>
F;3,92/Binding site: carbohydrate (Asn) #status predicted

Query Match 11.5%; Score 330; DB 2; Length 516;

Best Local Similarity 24.7%; Pred. No. 2.2e-17;

Matches 132; Conservative 93; Mismatches 197; Indels 112; Gaps 21;

Qy 4 RLWVSLRLHRAQLLLVNLTLTFGLVCLAAITVYPPLLLEVGVBEKFTWVLGIPVL 63

Db 29 KLM-----KIIWVASIAAGVQFGWALQSLTTPYVQLLGIHHTAAIYIWLCPIS 78

Qy 64 GLVCPVLLGSADHWGRGYRRRPFIFWALSGLTLLSLFLPRAGWLAGL-----LCPD 116

Db 79 GMIVQPVGVYSDNCSSRRFRRRPFIAAGALVMTAVFLI---GFAADLGHASGDTLGK 135

Qy 117 PRPLELAILGLVGLDFCGQVCFPTLEALLSDFLFRDPD--HCRQAYSVYAFMISLGGCLG 175

Db 136 FKPRALAVFVGVFWILDVANNMLOGPCRALLDLGGKGRMRTANAFSFFMAVGNILG 195

Qy 176 YLLPAIDW-----DTSALAPYLTQBECLF--GLTLTFLTCVAATLVVAEE--NALGP 225

Db 196 YAAGSYSHLFKVPFESKTRACDMYCANLKSFCFIFALLSLTTLTALTVRENEPEKDE 255

Qy 226 TPAEGLSAPSLSPHCCPCARLAFRNLCALLPRHLQCCRMPTLRRLFVABLCSWMAL 285

Db 256 QEIDEKLAGAG-----KSKVPF--FGEIFGALKEL-----PRPMWILLVTCINWIAW 301

Qy 286 MTFTLYTDFVGEGLYQGVPRAEPTGTEARRH--YDEGVRMGSGLFLQCAISLVFSLVMDRL 345

Db 302 FFFFLYDTDMAKEVFGG---QVGD--ARLYDLGVRAAGMGLLQSVVLGFWMSLGVFL 355

Qy 346 VORFGRTRAVYLASVAAPFVAAGATCLSHSVAV-----TASAALTGF----- 388

Db 356 GKKIG--GAKRLWGLNLFVL---AICLAMTILVTKMAEKSRQHPDAGTLPAGTGPVIGKAL 411

Qy 389 --PSALQI-----LPYTLASLYHREK-----QVFLPKYRGDTGGASSED 426

Db 412 LFLAALGIPLANTFSTIPFALASIFSSNRGSGGGLSLGVNLALIVPQMLVSLVGGWDDL 471

Qy 427 LMTSFLPGKPGAPFPNGHVGAAGSG-----LPPPPPACGASACDVSVRVVG 475

Db 472 FCGGNLPG-----FVVGAVAAASAVLALTMLPSPA-----DAKPAVAMG 512

RESULT 4

G84441

probable sucrose-proton symporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C/Accession: G84441
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; PMID:20083487; PMID:10617197
A/Accession: G84441
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-594 <STO>
A/Cross-references: GB:AE002093; NID:G3461813; PIDN:AAC32907.1; GSPDB:GN00139
C/Genetics:
A/Gene: At2g02860
A/Map position: 2
C/Superfamily: common tobacco sucrose transport protein

Query Match 11.5%; Score 328.5; DB 2; Length 594;
Best Local Similarity 24.1%; Pred. No. 3.3e-17;
Matches 120; Conservative 72; Mismatches 174; Indels 131; Gaps 18;

QY 22 NLLTGLEVCVLAAGITY-----VPLLLEVGVEEKEFTWVLGIGPVGLVLCVPLLGS 73
DB 59 SLVTLVLSCTVAGVQFGWALQSLTPIYQTLGISHAFSSFTWGLGPTGLVWQPFVGI 118
QY 74 ASDHWGRYGRRRPFTWALSIGLISLFIIPRAGWLA--GLLCPD-----PRPLE 121
DB 119 WSDKCTSKYGRRRPFTWALSIGLISLFIIPRAGWLA--GLLCPD-----PRPLE 175
QY 122 LALLIGVGLLPCGQVCTFPLEALLSDLFDPDCHCRQAYSVYAFMISGGCLGVLPAI 181
DB 176 AVVFIIGFWLLDANNVQGPALADL--SGPDQRNTANAVFLMMAIGNILGFSAGAS 234
QY 182 -----DWDTSALAPYLGTQECFLG-----LLFLFUT--CVAATLVVAEAAALGPTEP 228
DB 235 GKQEW-----FPFL--TSRACCAACGNLKAFLAVVFLTCTLTVTIYPAKEIPFTSNKP 288
QY 229 AE-GLSAPSLSPHCPCRRARAFRLN-----LSKGLHSLKNGTANGIKYRVERDTEQFNSENEHQ 340
DB 289 TRIQDSAPLLDD-----GALLPRHLQCCMRPTRLRLFAELCSWMALMTFTLYDFVGEGLYQVP 305
QY 341 DETYVDGFGSVLNLTLRLPAPMHVSLVIMWALTWLSWFFFLFDITDMGREVYHGD 400
QY 306 RAEPQTEARRHYDEGRVMSGLGFLQCAISLVFSLVMDRLVORFQTRAYV--LASVAAPV 364
DB 401 TGD--SLHMEYDQGVREGALGLLNSVVLGISTFLIEPCQRMGARVVMALSNFTVAC 458
QY 365 AGATCLSL-----HSVAVVTASALTGFTFSALQILPYTLASLY 403
DB 459 MAGTAVISLMSLSDKNGIEYIMRNETTRTAIV--FALLGFLAITYSVPFVSVA-- 514
QY 404 HREKQVFLPKYRGDTGG 420
DB 515 -----EVTDSGG 522

RESULT 5
S28052
sucrose transport protein - spinach
N/Alternate names: sucrose carrier protein; sucrose permease
C/Species: Spinacia oleracea (spinach)
C/Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C/Accession: S28052
R/Riesmeier, J.W.; Willmitzer, L.; Frommer, W.B.
EMBO J. 11, 4705-4713, 1992
A/Title: Isolation and characterization of a sucrose carrier cDNA from spinach by functi
A/Reference number: S28052; PMID:93099843; PMID:1464305
A/Accession: S28052
A/Molecule type: mRNA
A/Residues: 1-525 <RIE>

A/Cross-references: EMBL:X67125; NID:G21318; PIDN:CRA47604.1; PID:G21319
C/Superfamily: common tobacco sucrose transport protein
C/Keywords: transmembrane protein

Query Match 11.3%; Score 324.5; DB 2; Length 525;
Best Local Similarity 23.2%; Pred. No. 5.8e-17;
Matches 129; Conservative 91; Mismatches 196; Indels 141; Gaps 20;

QY 23 LUTFGLEVCLAAAGITY-----VPLLLEVGVEEKEFTWVLGIGPVGLVLCVPLLGS 74
DB 35 LKUGLNVASVAGVQFGWALQSLTPIYQTLGISHAFSSFTWGLGPTGLVWQPFVGI 94
QY 75 SDHWGRYGRRRPFTWALSIGLISLFIIPRAGWLA--GLLCPD-----RPLELALLIL 127
DB 95 SDRCTSRGRRRPFIAAGAAVAVAGLI--GFAADIGAASGDPPTGNVAKPRAIAV 151
QY 128 GVGILLDFGQVCTFPLEALLSDLF--DPDCHCRQAYSVYAFMISGGCLGVLPAID---- 182
DB 152 GFWILLDVANNITLQGPCRALLADMAAGSQTKTRYANAFSFFFMALGNIGGYAAGSYRLYT 211
QY 183 ---WDTSAAPYLGTQECFLGILT--LIFLTCVAAATLVAAEAAALGPTEPAEGLSAPSL 237
DB 212 VPPFTKTRACDVYCANLSCFPFISITLLVILILALSVMKQYITIDEIQEEDLNKRN 271
QY 238 SPHCPCRRARAFRLNIGALLPRHLQCCMRPTRLRLFAELCSWMALMTFTLYDFV 297
DB 272 SSGC---ARLPF--FGQLIGALKDL---PKPMLILLVLTALNWIAPFLLFDITDMG 321
QY 298 EGYGVPRAPRGTEARRHYDEGRVMSGLGFLQCAISLVFSLVMDRLVORFG--TRAVY 355
DB 322 KEVYGGT-----VOEGKLYDQGVHAGALGLMINSVILGVMSLSIEGLARMVGGAKLWG 375
QY 356 LASVAAPVAAACATCLSHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKYR 415
DB 376 IVNIIL-----AVCLMTV--LVTKSAE-----HFRDSSHIM----- 405
QY 416 GDTGASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVG 475
DB 406 -----GSVAPPPPPA---GVKGGALAFVILG 429
QY 476 EPTARV--VP-----GRGICLDLALDSAFLLSQV-----AP 506
DB 430 IPLAITFSIPFALASIFSAGSGGSLGVLNLAIVVQPMPFVSVTSGFWDAMFGGNLP 489
QY 507 SLFWGSIYQLQSQTAY 523
DB 490 APVVGAAVATAASVLSF 506

RESULT 6

S38196
sucrose transport protein SUC2 - Arabidopsis thaliana
N/Alternate names: sucrose-proton symporter SUC2
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000
C/Accession: S38196; T00773
R/Sauer, N.K.
submitted to the EMBL Data Library, October 1993
A/Description: SUC1 and SUC2: two sucrose transporters from Arabidopsis thaliana.
A/Reference number: S38196
A/Accession: S38196
A/Molecule type: mRNA
A/Residues: 1-512 <SAU>
A/Cross-references: EMBL:X75382; NID:G407091; PID:G407092
R/Vyotskaya, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Oji, O.; Kwan, A.; Liu, S.; Li,
rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N
submitted to the EMBL Data Library, July 1998
A/Description: Arabidopsis thaliana chromosome 1 BAC T22J18 sequence, complete sequence.
A/Reference number: Z14202
A/Accession: T00773
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-490, 'A', 492-512 <VYS>

A;Cross-references: EMBL:AC003979; NID:g3172156; GSPDB:GN00059; ATSP:T22J1
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: SUC2; ATSP:T22J18.12
A;Map position: 1
A;Introns: 419/3; 441/1; 455/3
C;Superfamily: common tobacco sucrose transport protein
C;Keywords: sugar transport

Query Match 11.3%; Score 323.5; DB 2; Length 512;
Best Local Similarity 24.1%; Pred. No. 6.7e-17;
Matches 123; Conservative 92; Mismatches 195; Indels 101; Gaps 18;

QY 14 RKAQLLVNLLTFGLVCLAAAGITVYVPPILLLEVGVEEKMVMVLGIPVGLVCPVLLGS 73
DB 28 RLKRTISVSSIAAGVQFGWALQSLTTPYVQLLGIPIHKWASLIWLCGIPSGMLVQPIVGY 87
QY 74 ASDHWRGRCRRRPFPIWALSIGLILSLFLIPRAGWLAGL-----LCPDPRPLEALLI 126
DB 88 HSDRCTSRFGRRRPFPIVAGAGLVTVAVFLI---GYAADIGHSGMDQLDKPKTKTAIPA 144
QY 127 LGVGLDFCGQVCFPTPLEALLSLFR-DPDHCRQAVSVYAFMISLGGCIGY-----LL 178
DB 145 LGFWILDVANNTLQGPCRAFLADLSAGNAKTRTANAFFSFFMAVGNVLGYAAGSYRNLY 204
QY 179 PAIDWD-TSALAPYLGTQBECLGLLTLFLTCVAATLLVABEALGPTPEAGLSAPSL 237
DB 205 KVPFTWTESCDLYCANLTKTFFLSITLL-LIVTFVSLCYVKEKPTPEPTADGKA---- 259
QY 238 SPHCCPCRARLAFRNIGALLPRHLQCCRMPTRLRLFVAELCSWMALMTFTLYTDFVG 297
DB 260 -----SNVPP--FGEIFGAFKEL-----KRPMMMLLIVTALNWIAPFLLFDTDWNG 305
QY 298 EGYQGVPRAPGTEARRHYDEGVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLA 357
DB 306 REVYGNSDATATAASKLYNDGVRAGALGLMLNAIVLGFMSLGVEWIGRKLK-GAKRLW 364
QY 358 SVAAPFAAGATCLSHSVAV-----VTASA-----ALTG-----FT 388
DB 365 GIWNFIL---AICLAMTVVTVTKQENHRDHGAKTGPNGVNTAGALTTLFAILGIPQAIT 421
QY 389 FSALQILPYTLASLYHREK-----QVFLPKYRGDTGGASSEDLSMTSLPLGP 435
DB 422 FS-----IPFALASIFSTNSGAGQGLSLGVNLAIIVPQWVISVGGGPFDELFGGNIPA- 476
QY 436 KPGAPPNGHVGGGSGLL-----PPPPA 459
DB 477 -----FVLGAIAAAVSGVLATVLPSPPPDA 502

RESULT 7
G86360
probable sucrose-proton symporter SUC2 protein T22J18.12 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Nov-2001
C;Accession: G86360
R;Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzbberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G86360
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-512 <STO>
A;Cross-references: GB:AE005172; NID:g3287687; PIDN:AAC25515.1; GSPDB:GN00141
C;Genetics:

A;Map position: 1
C;Superfamily: common tobacco sucrose transport protein

Query Match 11.3%; Score 323.5; DB 2; Length 512;
Best Local Similarity 24.1%; Pred. No. 6.7e-17;
Matches 123; Conservative 92; Mismatches 195; Indels 101; Gaps 18;

QY 14 RKAQLLVNLLTFGLVCLAAAGITVYVPPILLLEVGVEEKMVMVLGIPVGLVCPVLLGS 73
DB 28 RLKRTISVSSIAAGVQFGWALQSLTTPYVQLLGIPIHKWASLIWLCGIPSGMLVQPIVGY 87
QY 74 ASDHWRGRCRRRPFPIWALSIGLILSLFLIPRAGWLAGL-----LCPDPRPLEALLI 126
DB 88 HSDRCTSRFGRRRPFPIVAGAGLVTVAVFLI---GYAADIGHSGMDQLDKPKTKTAIPA 144
QY 127 LGVGLDFCGQVCFPTPLEALLSLFR-DPDHCRQAVSVYAFMISLGGCIGY-----LL 178
DB 145 LGFWILDVANNTLQGPCRAFLADLSAGNAKTRTANAFFSFFMAVGNVLGYAAGSYRNLY 204
QY 179 PAIDWD-TSALAPYLGTQBECLGLLTLFLTCVAATLLVABEALGPTPEAGLSAPSL 237
DB 205 KVPFTWTESCDLYCANLTKTFFLSITLL-LIVTFVSLCYVKEKPTPEPTADGKA---- 259
QY 238 SPHCCPCRARLAFRNIGALLPRHLQCCRMPTRLRLFVAELCSWMALMTFTLYTDFVG 297
DB 260 -----SNVPP--FGEIFGAFKEL-----KRPMMMLLIVTALNWIAPFLLFDTDWNG 305
QY 298 EGYQGVPRAPGTEARRHYDEGVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLA 357
DB 306 REVYGNSDATATAASKLYNDGVRAGALGLMLNAIVLGFMSLGVEWIGRKLK-GAKRLW 364
QY 358 SVAAPFAAGATCLSHSVAV-----VTASA-----ALTG-----FT 388
DB 365 GIWNFIL---AICLAMTVVTVTKQENHRDHGAKTGPNGVNTAGALTTLFAILGIPQAIT 421
QY 389 FSALQILPYTLASLYHREK-----QVFLPKYRGDTGGASSEDLSMTSLPLGP 435
DB 422 FS-----IPFALASIFSTNSGAGQGLSLGVNLAIIVPQWVISVGGGPFDELFGGNIPA- 476
QY 436 KPGAPPNGHVGGGSGLL-----PPPPA 459
DB 477 -----FVLGAIAAAVSGVLATVLPSPPPDA 502

RESULT 8
S48789
sucrose transport protein - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 26-May-2000
C;Accession: S48789
R;Buerkle, X.Y.Z.; Frommer, W.B.
submitted to the EMBL Data Library, October 1994
A;Reference number: S48789
A;Accession: S48789
A;Molecule type: mRNA
A;Status: preliminary
A;Residues: 1-507 <SUE>
A;Cross-references: EMBL:X82276; NID:g575350; PID:g575351
C;Superfamily: common tobacco sucrose transport protein

Query Match 11.1%; Score 317.5; DB 2; Length 507;
Best Local Similarity 24.3%; Pred. No. 1.9e-16;
Matches 125; Conservative 87; Mismatches 203; Indels 99; Gaps 18;

QY 4 RLWVSLRLHRRKAQLLVNLLTFGLVCLAAAGITVYVPPILLLEVGVEEKMVMVLGIPVGL 63
DB 23 KLV-----KIIMVASIAAGVQFGWALQSLTTPYVQLLGIPIHKFASFWLWCGPIS 72
QY 64 GLVCPVLLGSASDHWGRYGRRRPFTWALSIGLILSLFLIPRAGWLAGLCPD-----R 118
DB 73 GMIVQPVWGYSDNCSSRFRRGRFTAAAGAAVTVIAVFLIGFAADL-GHATGDLGKGSK 131
QY 119 PLEALLILGVLLDFCGQVCFPTPLEALLSLFR-DPDHCRQAVSVYAFMISLGGCIGYLL 178

```
Db 132 PRAIAVVFVGVFWLIDVANNMLOGPCRALLDLGGKARMTSNAPFFSFFMAVGNVLGYAA 191
QY 179 PAID-----WDTALAPVLTQBECLF-GLLTLFLTCVAATLLVAEEAALGPTEPA 229
Db 192 GSVSRLLCKIPFPFKTACDIYCANLKSCEFFIAVFLLSLTLALT--VRENELPEKDEH 249
QY 230 EGLSAPSLPHCCPCRARLARFNALGALLPRHLQCCRMPTLRRLFVAELCSMMALMTFT 289
Db 250 E-----IDEKAGARKSVFF--FGEIFGALKDL-----PRPMWILLVLTSLNIARPPFP 297
QY 290 LFTYDFVGEGLYQGVPAEPTGTEARRHYDEGVNMGSLGLFLOCAISLVFLVMDRLVQRF 349
Db 298 LYTDWMAKEVYGG--KVGDG----RLYDLGVHAGALGLLLNSVFLGFMLSVSEFLGKKI 351
QY 350 GTRAVVLASVAAPFVAAGATCLSHSVAV-----TASALGTGFTSALQI-- 394
Db 352 G-GVKRLWGLNLFV---AVCMALTIVLVTKMAEKSRQYDAGHTMPTSGVKIGALTFLFA 407
QY 395 -----LPYTLASLYHREK-----QVFLPKYRGDTGGASSEDLSMTS 430
Db 408 VLGIPLAVTFSPFALASIFSSNAGSGQLSLGVNLAIIVPQMLYSIAGGPWDDLFGGG 467
QY 431 FLPGPKGAPFNGHVGAGSG-----LLPPPPA 459
Db 468 NLPG-----FIVGAVAAAAGSILALTMLPSPPA 495

RESULT 9
S43142
sucrose transport protein - castor bean
N:Alternate names: sucrose carrier
C:Species: Ricinus communis (castor bean)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 26-May-2000
C:Accession: S43142
R:Weig, A.; Komor, E.
submitted to the EMBL Data Library, March 1994
A:Description: A sucrose carrier from Ricinus communis.
A:Reference number: S43142
A:Accession: S43142
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-533 <WEI>
A:Cross-references: EMBL:Z31561; NID:G468561; PID:G468562
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.9%; Score 311; DB 2; Length 533;
Best Local Similarity 23.3%; Pred. No. 6.2e-16;
Matches 134; Conservative 95; Mismatches 225; Indels 120; Gaps 19;

QY 17 QLLLVNLLTFGLEVCCLAAGITYVPPLLELVGVEEKFTWVLGIGPVGLVCVPLLGASD 76
Db 37 KVMVVASIAGIQFGWALQSLTTPVQLLIGIPHTWAAFICGLPGTSGMLVQPIVGYHSD 96
QY 77 HWRGYRRPFTWALSGLTLLSLFLIPRA---GWLAG-LLCPDPRPLELALLILGVGLL 132
Db 97 RCTSGRRPFTIASGAAVFAIVFLGYAADLGHLSGSLDKSPKTRAIAPFVGVFWIL 156
QY 133 DFCGVCFPTLEALLSDLF-RDPDHCROQYSVYAFMISLGGCLGYLLPAID-----W 183
Db 157 DVANNMLOGPCRALLDLGSQKTRTANALFPFMAVGNVLGYAAGAVTHLYKLPFT 216
QY 184 DTSALAPYLTQBECLFGLTLFLTCVAATLLVAEEAALGPTEPAEGL-----SAP 235
Db 217 KTTACDVYCANLKSCEFFIIVLLLSLTIVLALSIVYKEK----PWSPPQAVDNAEDDTASQA 272
QY 236 SLSPHCCPCRARL--AFRNIGALLPRHLQCCRMPTLRRLFVAELCSMMALMTFTLYT 293
Db 273 SSSAQPMPFFGELGAFKNL-----KEPMWILLVLTCLNIARPPFPFLFT 318
QY 294 DFWGEGLYQGVPAEPTGTEARRHYDEGVNMGSLGLFLOCAISLVFLVMDRLVQRFTRA 353
Db 319 DWMGREVYGG--DSSGSABQLKLYDRGVAGALGLMLNSVLTGFTSLGVEVLARGVG-GV 375
```

```
QY 354 VYLASVAAPFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPK 413
Db 376 KRLWGINVFL---AVCLAMTV-LVTQAESEST-----RR 405
QY 414 YRGDTGCGASEDSLMTSFLPGPKGAPFPNGHVGAGSGLLPPPPALCGASACDVSRVV 473
Db 406 FATVSGGAK-----VP-----LPPP---SGVKAGALALFAV 433
QY 474 VGEPTEARV-VP-----GRCICLDLALDSAPLLSQVAPSLFWGSIVOL--SOS 519
Db 434 MGVPQAITVISIPALASIFSNSTSGAGGSLGLVNLISIVIPQMIVSVAAGPMDALFGGN 493
QY 520 VTAYVMSAAGLGLVAIYFATQVVFQKSDLAITYSA 553
Db 494 LPAFVGAVALASGIFALTMLPSPQDMPSAKA 527

RESULT 10
T14339
sucrose-proton transport protein - carrot
N:Alternate names: sucrose/H+ symporter protein
C:Species: Daucus carota (carrot)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14339
R:Shakya, R.; Sturm, A.
plant Physiol. 118, 1473-1480, 1998
A:Title: Characterization of source- and sink-specific sucrose/H+ symporters from carrot
A:Reference number: Z17991; MUID:99063785; PMID:9847123
A:Accession: T14339
A:Molecule type: mRNA
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-501 <SHA>
A:Cross-references: EMBL:Y16766; NID:G2969886; PIDN:CAA76367.1; PID:G2969887
A:Experimental source: cultivar Namtaise; leaf
C:Genetics:
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.7%; Score 305.5; DB 2; Length 501;
Best Local Similarity 21.6%; Pred. No. 1.5e-15;
Matches 124; Conservative 91; Mismatches 213; Indels 145; Gaps 14;

QY 18 LLLVNLTFGLEVCCLAAGITYVPPLLELVGVEEKFTWVLGIGPVGLVCVPLLGASDH 77
Db 30 LURVASVACGIGFQWALQSLTTPVQELGIPHANSSIIWLCGPLSGLLVQPIVGHMSDQ 89
QY 78 WRGYYRRPFTWALSGLTLLSLFLIPRAGWLAGLL--CPDPRPLELALLILGVGLLDFC 135
Db 90 CTSKYRRPFTVAGGTAILAVIIIAHSADIGLLGDDTADNKTMAIVAFVIGFWILDVA 149
QY 136 QGVCFPTLEALLSDLF-RDPDHCROQYSVYAFMISLGGCLGYLLPAIDWDTSALAPYLT 194
Db 150 NNMTQCPCRALLADLTGNDARETRVANAYFSLFMAIGNVLGY-----ATGAYSGW 199
QY 195 QBECLFGLTLFLTC-----VAATLLVAEEAAL-----GPTPEAEG 232
Db 200 YKVFPSFTSCTINCANKSAFYDIIITITVISIAKERPRISSQDGPQFSEDGT 259
QY 233 SAPSLSPHCCPCRARLARFNIGALLPRHLQCCRMPTLRRLFVAELCSMMALMTFTLY 292
Db 260 AQ---SGHI-----EEAFWLWELFGTFRLLPGSVVWILLVLTCLNIARPPFP 305
QY 293 TDFVGEGLYQGVPAEPTGTEARRHYDEGVNMGSLGLFLOCAISLVFLVMDRLVQRFTR 352
Db 306 TDWGREYVYGGF-----NQQSYSQDGVNMGAFGLMNSVVLGITSVLMEKLCRIWGS 359
QY 353 AYLASVAAPFVAAGATCLSHSVAVVTASAAALGTFTESALQILPYTLASLYHREKQVFLP 412
Db 360 FW-----GLSNILMTTCFFAMLLITFIANKMDY----- 388
QY 413 KYRGDTGGASSEDLSMTSFLPGPKGAPFPNGHVGAGSGLLPPPPALCGASACDVSRVV 472
```

Db 389 -----GTNPPN-----GIVISALIVEA 406
QY 473 VVGPEARV-VP-----GRGICLDLAIDSAFLSQVAPSLFMSIVQL-----516
Db 407 ILGIPLAITYVPYALVSTRISLGLGQSLMGVNLAIWVQIVSLGSGWDQLFGGG 466
QY 517 -SQSVTAYMVSAAGLGLVAIYFATQVVDKSDL 548
Db 467 NSPAFVVAALSAPAAGLIALIAIRRRPRVDKSL 499

RESULT 11
S48788
sucrose transport protein - tomato (fragment)
C:Species: Lycopersicon esculentum (tomato)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jun-2000
R:Buerkle, X.Y.Z.; Frommer, W.B.
submitted to the EMBL Data Library, October 1994
A:Reference number: S48787
A:Accession: S48788
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-428 <RUE>
A:Cross-references: EMBL:X82275
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.4%; Score 297.5; DB 2; Length 428;
Best Local Similarity 25.8%; Pred. No. 5.1e-15;
Matches 114; Conservative 82; Mismatches 163; Indels 83; Gaps 18;
QY 4 RLWVRLRRKQAQLLVNLTFLGLVCLAAAGITYVPPILLLEVGVBEKMTVMVLGIGPVYL 63
Db 25 KLV-----KIVVASIAGVQFGWALQSLTTPVQVLLGIPHRFASFIWLCGPIS 74
QY 64 GLVCPVLLGASDPMRGYRRRPFTWALSGLILLSLFLIPRAGWLAGLLCPDP-----R 118
Db 75 GMTVQPVGVYSDNCSSRFRGRRPFTIAGAAVTIAVFLIGFAADL-GHAGSDPLGKSK 133
QY 119 PLELALLIIGVLLDFCGQVCTPPLLEALLSDLRDPD-HCRQAYSVYAFMISLGGCLGY- 176
Db 134 PRAIAVVFVGFWDLVANNMLOGPCRALLDLADLSGGSKGKRTANAFSPFKAVGNILGYA 193
QY 177 -----LLPAIDWD-TSALAPYLGTQECIF-GLTLTILFTCVAAATLL-----VAEEAALG 224
Db 194 AGSYSLFKVFPFSKTKACDMYRANKSCFFIAIFLLSLTTLALTLVRENELPEKEELE 253
QY 225 PTEPAEGLSAPSLSPHCCPCRLARLAFRNGLALLPRHLQCCMRPTRLRLFVAELCSWMA 284
Db 254 IDEKLSGAG-----KSKVPF--FGEIFGALKDL-----PRPMWILLTLCLNWIA 296
QY 285 LMTFTLFTYDFVGEGLYQGVPRAPGTEARRHYDEGRMGSLGLFLOCAISLVFSLVMDR 344
Db 297 WFPFFLYDTDMWAKEVFGG---QVGD--AKLYDLGVRAGALGLLLQSVVLGFMSLGVF 350
QY 345 LVQRFCTRAVLASVAAPFVAAGATCLSHSAVV-----TASAAATGFT-----388
Db 351 LGKKIG-GAKRLWGIINFLV--AICLAWTILVTKMAEKSRGHAAGTLMGPTPGVKIGA 406
QY 389 ---FSALQI-----LPVTLA 400
Db 407 LLLFALGIPLAFTFSIPPALA 428

RESULT 12
S38657
sucrose transport protein ptpl - common plantain
N:Alternate names: sucrose transporter ptpl
C:Species: Plantago major (common plantain)
C:Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 26-May-2000
C:Accession: S38657
R:Gahrtz, M.; Stolz, J.; Sauer, N.
submitted to the EMBL Data Library, November 1993

A:Reference number: S38657
A:Accession: S38657
A:Molecule type: mRNA
A:Residues: 1-510 <GAH>
A:Cross-references: EMBL:X75764; NID:g415987; PID:CAA53390.1; PID:g415988
C:Genetics:
A:Gene: ptpl
A:Superfamily: common tobacco sucrose transport protein
C:Keywords: sugar transport

Query Match 10.3%; Score 294; DB 2; Length 510;
Best Local Similarity 24.7%; Pred. No. 1.1e-14;
Matches 126; Conservative 77; Mismatches 199; Indels 108; Gaps 19;
QY 18 LLLVNLTLFGLVCLAAAGITYVPPILLLEVGVBEKMTVMVLGIGPVLLGASDPM 77
Db 29 IFLVAIAAGVQFGWALQSLTTPVQVLLGIPHKWASIVLWCGPISGMIVQVGVFSDN 88
QY 78 WRGYYRRRPFTWALSGLIL-LSLFLIPRAGWLAGL-----LCPDPRLELALLILGV 129
Db 89 CTSRFRGRRPFI-AAGAGLVGVAVVLI---GPAADLGHAGSDSLGDLKPRAIGVFVFGF 144
QY 130 GLLDFCGQVCTPPLLEALLSDLF-RPDHCRQAYSVYAFMISLGGCLGYLLPAID-----182
Db 145 WILDVANNMLOGPCRALLDLADLSGGNTKMANANSFFPFMAVGNVLGYAAGSYRMYKVF 204
QY 183 --WDTSAAPYLGTQECIF-LFGLTLTILFTCVAAATLL-----VAEEAALGTEPAEGLSA 234
Db 205 PRSKTKACDIYCANLKSCFFIISITLITLTLALSIIVREKRVAREOV---TAAKGPKI 261
QY 235 PSLSPHCCPCRLARLAFRNGLALLPRHLQCCMRPTRLRLFVAELCSMMALMTFTLYTD 294
Db 262 P-----VPELFGALKDLPRPMWVLLLVTLNLTALNTIAWFGFLFDTD 301
QY 295 FVGEGLYQGVPRAPGTEARRHYDEGRMGSLGLFLOCAISLVFSLVMDRIVQRFG-----350
Db 302 WMGREVY-GETOOHKAPELAVIYNKGVSAAGALGLMLNSIVLGFASIGVQVMARALGGVKR 360
QY 351 -----TRAVLA-SVAAPFVAAGATCLSHSAVVVTAALTG--FTFSALQI-----394
Db 361 LMGVNFILAICTMTIVITVKASHRPYSNGVLQTPSSVKIGALVVFESALGIPLAITF 420
QY 395 -LPYTLASLYHREKQVFLPKYRGDTGGASSEDLSL-----MTSFLPGPKPG-----438
Db 421 SVFPFALASIV-----STTTGSGQSLGLVNLAIVIPQMIVSVASGPMWDMFEG 468
QY 439 ---APFPNGHVAGSGG-----LLPPPPA 459
Db 469 GGNLPAFVVGAVAAASGIFAFMTLPSPPA 498

RESULT 13
S51114
sucrose-proton symporter - beet
C:Species: Beta vulgaris (beet)
C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 26-May-2000
C:Accession: S51114
R:Westram, A.; Eckhardt, U.; Frommer, W.B.; Riesmeier, J.W.
submitted to the EMBL Data Library, January 1995
A:Description: Sequence of a sugar beet sucrose transporter cDNA.
A:Reference number: S51114
A:Accession: S51114
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-523 <WES>
A:Cross-references: EMBL:X83850; NID:g633171; PID:g633172
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.3%; Score 294; DB 2; Length 523;
Best Local Similarity 24.3%; Pred. No. 1.2e-14;
Matches 125; Conservative 78; Mismatches 199; Indels 112; Gaps 20;
QY 17 QLLVNLTLFGLVCLAAAGITYVPPILLLEVGVBEKMTVMVLGIGPVLLGASD 76

Db 38 KLALVASIAAGVQFGWALQSLTTPYVQLLGIPTHTWAPYIWLCCGPISGMIVQPTVGYSD 97
Qy 77 HWRGRRRRPFIWALSIGLLSLFLPRAGWLA--GLLCPD-----PRPLEALLILGV 129
Db 98 RCTSKFRRRPFPI---AVGATLVGFVSLVGLFAADI GHATGDPNGNVKPRATIAVVFVG 154
Qy 130 GLLDPCGQVCFPLEALLSLDFRDP--HCRQAVSVYAFMISLGGCLGY-----LLP 179
Db 155 WILDVANNTLQGPCRALADWAGSQAKTRYANAFSPFPMALGNIGYAGSYGLYTVFP 214
Qy 180 AIDWTSALAPYLGTQBECLFGLTLFLICVAATLLVABEALGPTPEAGLSAPSLSP 239
Db 215 FT--HTRACDTCYCANLKSCEFSITLLIVLTILALSVRER---PFTLDEIQEENLKN 268
Qy 240 HCCPCRLARLAFNIGALLPRHLQCCRPRTLRFLVAELCSNMALMTFTLFTVDFGEG 299
Db 269 NTGGC--ARLFP--FGQLFGALKD-----PKPMLILLVTCNLINAWPFLFLFDMDMGE 321
Qy 300 LYQGVPRAEFGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFQ--TRAVILA 357
Db 322 VYGT-----VGBGKAYDGVHAGALGLMINSVVLGIMSLGIEKLARLVGGVRLNGIV 375
Qy 358 SVAAPFVAAGATCLSHSVAVVTASA-----ALTF 387
Db 376 NLIL-----AVCLAMTI-LVTKSAEHRATHRVPGAIGPPLPPPGVKGGAIFAVALGI 428
Qy 388 TFSALQILPYTLASLYHREK-----QVFLPKYRCDTGASSEDSLMT 429
Db 429 PLAITTSIPFALASIFSASSGSGQLSLGVNLINIAVVPQMFVSVTSGPW-----DAL-- 480
Qy 430 SFLPGPKPGAPFNGHVGAGSG-----LLPppp 458
Db 481 -FGGNLPA--FVUGVAATASAILSTLLPppp 511

RESULT 14

F96741
probable sucrose transport protein p17M19.4 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: F96741
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, X.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96741
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-512 <STO>
A;Cross-references: GB:AE005173; NID:G6978914; PIDN:AAF34306.1; GSPDB:GN00141
C;Genetics:
A;Gene: F17M19.4
A;Map position: 1
C;Superfamily: common tobacco sucrose transport protein

Query Match 10.2%; Score 292.5; DB 2; Length 512;
Best Local Similarity 23.8%; Pred. No. 1.5e-14;
Matches 120; Conservative 87; Mismatches 211; Indels 87; Gaps 18;

Qy 17 QLLLVNLLTFGLVCLAGITYVPPILLLEGVVEKEFTMWLIGVILGVLCVPLIGSASD 76
Db 33 KIISVASIAAGVQFGWALQSLTTPYQLLGIPTHTWAPYIWLCCGPISGMIVQPTVGYSD 92
Qy 77 HWRGRRRRPFIWALSIGLLSLFLPRAGWLA-----LCPDPRPLEALLILGV 129

Db 93 RCBSEFRRRPFIAGVALVAVSVFLI---GFAADMGHSHFGDKLENKVRTRAIILFTGF 149
Qy 130 GLLDPCGQVCFPLEALLSLDFR--DPDHCRQAVSVYAFMISLGGCLGY-----LL 178
Db 150 WFLDVANNTLQGPCRAFLADLAAGDAKTRVANACSFPMVGNVLGYAAGSVTNLHKMF 209
Qy 179 PAIDWTSALAPYLGTQBECLFGLTLFLICVAATLLVABEALGPTPEAGLSAPSL 238
Db 210 PFT--MTKACDIYCANLKTCPFLSITLLIVLTFSSLYVVKD---QWSPQDQKEEKS 263
Qy 239 PHCCPCRLARLAFNIGALLPRHLQCCRPRTLRFLVAELCSNMALMTFTLFTVDFGEG 298
Db 264 -----SLFF--FGEIFCAVR---HMKRPMVMLLIVTINWIAWFFFLYDTDMGR 309
Qy 299 GLYQGVPRAEFGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLA 358
Db 310 EVTGG--NSDGRSRSKKYDQGVQAGALGMFNSILLGPFVSLGVESIGRMKG--GAKRLMG 366
Qy 359 VAAFPVAAGATCLSHSVAVVTASA-----SAALTGFTFSALQIL----- 395
Db 367 CVNFILAI--LMTVLTAKSAEHRRETAGPLAGPSSGKAGVFSLFTVLGILAITYS 423
Qy 396 -PYTLASLYHREKQVFLPKYRCDTGASSEDSLMTSFLFPKP-----GAPFN---GHVG 447
Db 424 IFFALASIFSTNSGAGQGLSLGVNLINIAICIPQMIVSFSSGPLDAQFGGGLPSFVVGAIA 483
Qy 448 AGSGILL-----PPP--PALCGA 463
Db 484 AAVSGVLALTLPSPPPDAPNMSG 508

RESULT 15

T02982
probable sucrose transport protein - rice
C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C;Accession: T02982
R;Hirose, T.; Inaizumi, N.; Scofield, G.N.; Furbank, R.T.; Ohsugi, R.
Plant Cell Physiol. 38, 1389-1396, 1997
A;Title: cDNA cloning and tissue specific expression of a gene for sucrose transporter f1
A;Reference number: Z14809; MUID:98182940; PMID:9522469
A;Accession: T02982
A;Status: preliminary;
A;Molecule type: mRNA
A;Residues: 1-537 <HIR>
A;Cross-references: EMBL:D87819; NID:G2723470; PIDN:BAA24071.1; PID:G2723471
A;Experimental source: cultivar Nipponbare, leaf
C;Genetics:
A;Gene: SUT1
C;Superfamily: common tobacco sucrose transport protein

Query Match 10.2%; Score 292; DB 2; Length 537;
Best Local Similarity 26.7%; Pred. No. 1.7e-14;
Matches 115; Conservative 72; Mismatches 167; Indels 76; Gaps 19;

Qy 17 QLLLVNLLTFGLVCLAGITYVPPILLLEGVVEKEFTMWLIGVILGVLCVPLIGSASD 76
Db 51 RLILSGMVAGGVQYQWALQSLTTPYQTLGSHALTSPFMLCGPIAGWVQPCVGLYSD 110
Qy 77 HWRGRRRRPFIWALSIGLLSLFLI---PRAGWLAGLCPD-----PRPLEALLIL 127
Db 111 RCTSKWRRRPYILTGCVLICLVAVVIGFSDIGADYMGDTKEDCSVYHGRWHAAIVVL 170
Qy 128 GVLDDPCGQVCFPLEALLSLDF--RDPDHCRQAVSVYAFMISLGGCLGYLLPAID--- 182
Db 171 GFWLLDFSNNTVQGPALMADLSGRHGP---TANSIFCSMMAMGNILYSSGSTNNWH 227
Qy 183 -WDTSAAPYLGTQBE-----LFG--LLTILFLT-CVAATLLVABE-----AALGPT- 226
Db 228 KW-----FPFLKTRACCACANLKGAFVAVIFLSLCIVITLIFAKEVDFPKGNAAL--PTK 281
Qy 227 --EPAGELSAPSLSHCCPCRLARLAFNIGALLPRHLQCCRPRTLRFLVAELCSNWA 284

Db	282	SNEPAREGTG-----PLAVLKGRNLTGMPV-----LIVTGL-TWLS	320
Qy	285	LMTFTLFYTDVCEGLYQGVPR-AEPGTEARRHYDEGRVMSIGLFLQCAISLVFSLVMD	343
Db	321	WPFPIYDTDMGREIYHGDPKGTDPQIEA---FNOGVRAQAFGLLNSIVLGFSSFLIE	377
Qy	344	RLVQREGTTRAVYLASVAAPFVAAGATCL-----SHSVA--VVTASAALTGFTFSAIQ	393
Db	378	PMCRKVGPRVWVTSNFLVCIAAATAAISFWSLKDFHGTQVOKAITADKSIKAVCLVLEA	437
Qy	394	ILPYTLASLY	403
Db	438	FLGVPLAVLY	447

Search completed: December 3, 2003, 17:21:14
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 17:16:16 ; Search time 18 Seconds
(without alignments)
1444.765 Million cell updates/sec

Title: US-09-593-793a-113

Perfect score: 2861

Sequence: 1 MVQRLVSRLLRHRKAQLK.....AIYFATQVVFDSKLAKISA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	489.5	17.1	530	1 MATP_MOUSE	P58355 mus musculus
2	488.5	17.1	530	1 MATP_HUMAN	Q9umx9 homo sapien
3	324.5	11.3	525	1 STP_SPTOL	O03411 spinacia ol
4	235.5	8.2	553	1 SUTI_SCHPO	O14091 schizosacch
5	142	5.0	544	1 YD74_SYNY3	P74168 synecocyst
6	136	4.8	541	1 GT10_HUMAN	O95528 homo sapien
7	125.5	4.4	399	1 TCRI_ECOLI	P02982 escherichia
8	121.5	4.2	503	1 PUR8_STRLP	P42670 streptomyce
9	119.5	4.2	491	1 AMPG_ECOLI	P36670 escherichia
10	117	4.1	440	1 YHJE_ECOLI	P37643 escherichia
11	117	4.1	495	1 SMVA_SALTY	P37594 salmonella
12	111.5	3.9	640	1 Y051_MYCTU	Q10880 mycobacteri
13	111	3.9	473	1 PHDK_NOSCK	O24723 nocardioid
14	110.5	3.9	680	1 CALA_HUMAN	Q03692 homo sapien
15	110	3.8	368	1 GALT_HUMAN	O60755 homo sapien
16	110	3.8	606	1 NOOC_THETH	O56227 thermus the
17	109.5	3.8	654	1 SPH2_HUMAN	Q9nra0 homo sapien
18	109	3.8	477	1 YPUM_RHOCA	P26176 rhodobacter
19	107	3.7	419	1 CMLA_PSEAE	P32482 pseudomonas
20	105	3.7	448	1 YJ94_YEAST	P47159 saccharomyc
21	104	3.6	476	1 MELE_SALTY	P30878 salmonella
22	103	3.6	465	1 FTSW_MYCLE	O50186 mycobacteri
23	102.5	3.6	461	1 PUCC_RHOCA	P23462 rhodobacter
24	102	3.6	438	1 SHIA_ECOLI	P76350 escherichia
25	101.5	3.5	428	1 YXIO_BACSU	P42306 bacillus su
26	101.5	3.5	510	1 NANT_YERPE	Q8zch3 yersinia pe
27	101	3.5	437	1 BRAZ_PSEAE	P25185 pseudomonas
28	101	3.5	471	1 MELE_ENTAE	O07366 enterobacte
29	101	3.5	471	1 MELE_KLEPN	O02581 klebsiella
30	100	3.5	404	1 Y4XM_RHISN	P55705 rhizobium s
31	100	3.5	473	1 Y1HO_SALTY	Q911r4 salmonella
32	99	3.5	386	1 CV03_HUMAN	Q9y3p4 homo sapien
33	99	3.5	481	1 LMRA_STRLN	P46104 streptomyce

RESULT 1

ID	MATP_MOUSE	STANDARD;	PRT;	530 AA.
AC	P58355;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Membrane-associated transporter protein (AIM-1 protein) (Melanoma antigen AIM1) (Underwhite protein).			
DE	MATP OR AIM1 OR UW			
GN	Mus musculus (Mouse).			
OS	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=dy; TISSUE=Eye, Kidney, and Uterus;			
RX	MEDLINE=21372467; PubMed=11479596;			
RA	Fukamachi S., Shimada A., Shima A.;			
RT	"Mutations in the gene encoding B, a novel transporter protein, reduce melanin content in medaka."			
RL	Nat. Genet. 28:381-385(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND VARIANTS UW-DBR ASN-153 AND PRO-435..			
RX	MEDLINE=21473748; PubMed=11574907;			
RA	Newton J.M., Cohen-Barak O., Hagiwara N., Gardner J.M., Davissom M.T., King R.A., Brilliant M.H.;			
RT	"Mutations in the human orthologue of the mouse underwhite gene (uw) underlie a new form of oculocutaneous albinism, OCA4."			
RL	Am. J. Hum. Genet. 69:981-988(2001).			
CC	-!- FUNCTION: Melanocyte differentiation antigen. May transport substances required for melanin biosynthesis (By similarity).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein; melanosome (By similarity).			
CC	-!- TISSUE SPECIFICITY: Melanocytes, eyes, kidney and uterus.			
CC	-!- DISEASE: Defects in MATP are the cause of the UW-dbr phenotype that results in loss of nearly all pigmentation in the homozygous state.			
CC	-!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; AF360357; AAK81713.1; --			
DR	MGI; MGI:2153040; Matp.			
KW	Melanin biosynthesis; Transmembrane; Antigen; Glycoprotein; Vision; Disease mutation; Albinism.			
KW	DOMAIN 1 45 CYTOPLASMIC (POTENTIAL).			
FT	DOMAIN 46 66 1 (POTENTIAL).			
FT	DOMAIN 67 68 EXTRACELLULAR (POTENTIAL).			
FT	DOMAIN 69 89 2 (POTENTIAL).			

P33607 escherichia
O87656 salmonella
O79451 sorex ciner
P28568 gallus gall
Q9evn4 pseudomonas
P79393 bos taurus
O34245 wolfinella s
Q02846 homo sapien
Q79452 sorex hayde
O79969 sorex monti
O80019 sorex vagra
O88626 rattus norv

34 98.5 3.4 613 1 NUOL_ECOLI
35 98.5 3.4 685 1 FHUB_SALTY
36 98 3.4 379 1 CYB_SORCI
37 98 3.4 496 1 GTR3_CHICK
38 97.5 3.4 354 1 RNFD_PSEST
39 97.5 3.4 385 1 P12R_BOVIN
40 97.5 3.4 433 1 DCUA_WOLSU
41 97.5 3.4 1103 1 CYGD_HUMAN
42 97 3.4 336 1 CYB_SORHA
43 97 3.4 336 1 CYB_SORMO
44 97 3.4 336 1 CYB_SORVA
45 97 3.4 370 1 GAUT_RAT

ALIGNMENTS

FT	DOMAIN	90	105	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	106	126	3 (POTENTIAL).
FT	TRANSMEM	127	138	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	139	159	4 (POTENTIAL).
FT	TRANSMEM	160	184	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	185	205	5 (POTENTIAL).
FT	TRANSMEM	206	216	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	217	237	6 (POTENTIAL).
FT	TRANSMEM	238	318	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	319	339	7 (POTENTIAL).
FT	TRANSMEM	340	366	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	367	387	8 (POTENTIAL).
FT	TRANSMEM	388	398	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	399	419	9 (POTENTIAL).
FT	TRANSMEM	420	425	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	426	446	10 (POTENTIAL).
FT	TRANSMEM	447	477	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	478	498	11 (POTENTIAL).
FT	TRANSMEM	499	504	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	505	525	12 (POTENTIAL).
FT	DOMAIN	526	530	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	356	356	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	153	153	D -> N (IN UW-DBR).
FT	VARIANT	435	435	S -> P (IN UW-DBR).
SQ	SEQUENCE	530 AA;	57961 MW;	F4EDEAD07916D9FC CRC64;

Query Match
 Best Local Similarity 26.1%; Pred. No. 2,7e-28;
 Matches 146; Conservative 85; Mismatches 228; Indels 101; Gaps 12;

QY	17	QLLLVLLTGLVCLAAAGTYVPPILLLEGVVEKEFTWVLGPGVLGCVPLGCSAD	76
DB	34	RLVHNSMFGREFCYAVEAAYTTPVLLSVGLPKSLYVMVLLSPILQLQVVGASD	93
QY	77	HWRGRRPFTWALSGLILLSLFTIPRAGWLAGLLCPDPR-LEALILGVLGLD	133
DB	94	HCBARWRPPTLTALMMLGMALYNGDAVVSALVANFRQKLIWASITWGVVLPD	153
QY	134	FCQGVCTPLEALLSLDFRDPDHCRAQSVYAFMISLGGCLGYLLPAIDWDTSALAPYLG	193
DB	154	FSADFTDGPVKAYLFDVCSHQDK-EGKHLVHALFTGGALGYILGAIDWVHLDLGRLG	212
QY	194	TOECLFGLLTFLFCVAATLVAEEAL--GPTPE-----ASGLSAPSLPHCCPC	244
DB	213	TEFQVFFSALVLLICFTLHCISIPEAPURDAATDPPSQDPPQSSLSASGMHEY----	268
QY	245	RARLAFRLGA-----LLPRLHQLCCMRPRTLRRLFLVABLCSW	282
DB	269	GSIEKVKNGGADTEQPVQEMKKNKPSQSQRTWSMKSLLRALVNPMSHYRCLCVSHLIGW	328
QY	283	MALMTTFLTYDFVGBGLYGVPRAPFGTEARRHYDEGVPMGLGLFLQCAISLVFLVM	342
DB	329	TAFLSNMLFTDPMGQIVVHGDYGAHNSTEFILYERGVGCGWGLCINSVFSSVSYFC	388
QY	343	DRIVQRFGRTRAVLASVAAPPAAGATCLSHSVAVVTASAAALGFTFSAQLIIPYTLASL	402
DB	389	KAMVSIVGLKGLYFMGLVFLGLGTGFIPLFPNYSITLVLCSMFVMSSTLYVFPFLIAE	448
QY	403	YHREKQVFLPKYRGDTGGASSEDLSMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCG	462
DB	449	YHREEE-----KEKGQEA-----PGGPDNQR-----DCA	477
QY	463	ASACDVSVRVVGTEPTARVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIVOLSOVTA	522
DB	478	ALTCMVQL-----AQILVGGG-----LGLFLNMGASVVV	506
QY	523	YVVSAGLGLVAIYFATQVV	542
DB	507	VVTASAVSLIGCCFVALFV	526

RESULT 2
 MATP_HUMAN

ID	MATP_HUMAN	STANDARD;	PRT;	530 AA.
AC	Q9UMX9; Q9BTM3;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Membrane-associated transporter protein (AIM-1 protein) (Melanoma antigen AIM1).			
DE	MATP OR AIM1.			
GN	Homo sapiens (Human).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]_TaxID=9606;			
RP	SEQUENCE FROM N.A. (ISOFORM AIM-1A).			
RC	TISSUE=Melanoma;			
RX	MEDLINE=21115844; PubMed=11221837;			
RA	Harada M., Li Y.F., El-Gamil M., Rosenberg S.A., Robbins P.F.;			
RT	"Use of an in vitro immunoselected tumor line to identify shared melanoma antigens recognized by HLA-A*0201-restricted T cells.";			
RL	Cancer Res. 61:1089-1094 (2001).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM AIM-1B).			
RC	TISSUE=Skin;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richardson D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Whiting M., Madan A., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,			
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).			
RN	[3]			
RP	ALTERNATIVE SPLICING.			
RA	Ferro S.;			
RL	Unpublished observations (NOV-2001).			
RN	[4]			
RP	DISEASE AND VARIANT LEU-374.			
RX	MEDLINE=21473748; PubMed=11574907;			
RA	Newton J.M., Cohen-Barak O., Hagiwara N., Gardner J.M., Davison M.T.,			
RA	King R.A., Brilliant M.H.;			
RT	"Mutations in the human orthologue of the mouse underwhite gene (uw) underlie a new form of oculocutaneous albinism, OCA4.";			
RL	Am. J. Hum. Genet. 69:981-988 (2001).			
CC	-!- FUNCTION: Melanocyte differentiation antigen. May transport substances required for melanin biosynthesis (By similarity).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein; melanosome (By similarity).			
CC	-!- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=3;			
CC	Comment=Additional isoforms seem to exist;			
CC	Name=AIM-1a;			
CC	isoId=Q9UMX9-1; Sequence=Displayed;			
CC	Name=AIM-1b;			
CC	isoId=Q9UMX9-2; Sequence=VSP_006297, VSP_006298, VSP_006299;			
CC	Name=AIM-1c;			
CC	isoId=Q9UMX9-3; Sequence=VSP_006296;			
CC	-!- TISSUE SPECIFICITY: Expressed in most melanoma cell lines and melanocytes.			
CC	-!- DISEASE: Defects in MATP are the cause of oculocutaneous albinism type 4 (OCA4). OCA4 is characterized by hypopigmentation of skin,			


```

DR Pfam: PF00083; sugar_tr; 1.
DR TIGRfam: TIGR01301; GPH_sucrose; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; FALSE NEG.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
KW Transmembrane; Transport; Sugar transporter; Symport.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 58 1 (POTENTIAL).
FT TRANSMEM 72 92 2 (POTENTIAL).
FT TRANSMEM 107 127 3 (POTENTIAL).
FT TRANSMEM 145 165 4 (POTENTIAL).
FT TRANSMEM 184 204 5 (POTENTIAL).
FT TRANSMEM 230 250 6 (POTENTIAL).
FT TRANSMEM 295 315 7 (POTENTIAL).
FT TRANSMEM 338 358 8 (POTENTIAL).
FT TRANSMEM 373 393 9 (POTENTIAL).
FT TRANSMEM 422 442 10 (POTENTIAL).
FT TRANSMEM 455 475 11 (POTENTIAL).
FT TRANSMEM 488 508 12 (POTENTIAL).
FT DOMAIN 509 525 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 525 AA; 54992 MW; 018347A4D2C1CG6 CRC64;

Query Match 11.3%; Score 324.5; DB 1; Length 525;
Best Local Similarity 23.2%; Pred. No. 2.5e-16;
Matches 129; Conservative 91; Mismatches 196; Indels 141; Gaps 20;

QY 23 LLTFGLVCLAAGITY-----VPLLLEGVVEKEFTMWLGIGPVGLVLCVPLLGSA 74
Db LKGLGLVASVAAGVQFGWALQLSLTPYVQLLGPHTWAAIYWLCPGISGMIVQPLVGY 94
QY 75 SDHWRGVRGRRPFIWALSIGLLSLFLIPRAGWLA--GLLCDDP-----RPLELALLIL 127
Db SDRCTSRGRRRPFIAAGAAALVAVAGLI---GFAADIGAAAGDPTGNVAKPRAIYFVV 151
QY 128 GVGLLDFCGCVCTPLEALLSDLFR-DPDHCRQAVSYAFMISLGLGCLGYLLPAID- 182
Db GFWILDVANNTLQGPCRALLADMAAGSQTKTRYANAFPSFMALGNIGVYAAGSYRLYT 211
QY 183 -----WDTSAIAPYLGTOEBCFLGILT-LIPLTCAATLLVAEAAALGPTPEAGLSAPSL 237
Db VFPFTKTAACDVYCANLKSCFFISITLLIIVLTILALSVMKQRQITIDEIQEEDLKNRN 271
QY 238 SPHCCPCRLAFRNIGALLPRHLQCCRRPTRLRLFVAELCSWMAALMTFTLFYTDVFG 297
Db SSGC-----ARLPF--FGQLGALKDL-----PKWMLILLITLALNWTANWFFLLPDTDWG 321
QY 298 EGYLQGVPRAPGPTGARHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFQ--TRAVY 355
Db KEVYGGT-----VGEKLYDQGVHAGALGMINSVVLGVMSLSIEGLARMVGGAKRLWG 375
QY 356 LASVAAPPVAAGATCLSHSVAVVTASALTGFTFSALQILPYLASLYHREKQVFLPKYR 415
Db IVNII-----AVCLAMTV-LVTKSAE-----HFRDSSHIM----- 405
QY 416 GDTGGASSEDLSMTSLFPGPKGAPFNGHVAGGSGLLPPPALCGASCDVSVRVVWG 475
Db -----GSAVPPPPPA--GVKGGALAIYFAVLG 429
QY 476 EPTEARV-VP-----GRGICLDLAILDSAFLLSQV-----AP 506
Db IPLAITFISIPFALASIFASSGSGQGLSLGLVNLAIIVPQMFVSVTSGPDMAFGGGNLP 489
QY 507 SLFMGSIVQLSQSVTAY 523
Db AFVVGAAVATASAVLSF 506

RESULT 4
ID SUT1 SCHPO STANDARD; PRT; 553 AA.
AC O14031; Q9UUG0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

```

DE General alpha-glucoside permease.
GN SUT1 OR SPAC2F3.08.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RX Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feitwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinhardt R., Fournelle B.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fournelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Moreno S., Armstrong J., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Benito J.,
RA Cerutti L., Lowe T., McComb W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
[2]
RN SEQUENCE OF 35-166 FROM N.A.
RP STRAIN=968 h90;
RC MEDLINE=20233868; PubMed=10759889;
RX Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
RA Hirakawa Y.;
RT "Large-scale screening of intracellular protein localization in living
RT fission yeast cells by the use of a GFP-fusion genomic DNA library";
RL Genes Cells 5:169-190(2000).
[3]
RN IDENTIFICATION, AND FUNCTION.
RP MEDLINE=21065106; PubMed=11136464;
RX Reinders A., Ward J.M.;
RA "Functional characterization of the alpha-glucoside transporter SUT1p
RT from Schizosaccharomyces pombe, the first fungal homologue of plant
RT sucrose transporters";
RL Mol. Microbiol. 39:445-454(2001).
CC -!- FUNCTION: Responsible for the transport of maltose and sucrose
CC into the cell, with the concomitant export of a proton (Symport
CC system).
CC -!- PATHWAY: Maltose and sucrose metabolism.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z99165; CAB16264.1; -.
CC EMBL; AB027797; BAA87101.1; -.
CC PIR; T38541; T38541.
CC GenedB SPombe; SPAC2F3.08; -.
CC Transmembrane; Transport; Sugar transporter; Symport.

Db 109 GFAISLSMACCIYVSELV-GPQRQGVLSYELVAGITVIGLLSYALNYALAGTWPWGRHM 167
 QY 187 ---ALAPYLGTQBECLFGLLTLLFTCTVAATLLVAEEALGPTPEAGLSAPLSHCCP 243
 Db 168 FGWATAP-----AVLQSLSLFLP--AGT---DETATKDLIPQGGAPKLGFG--- 212
 QY 244 CRARLAFRNIGALLPRHLQJCCMPRTLRFLVAELCSNMALMTFTLYTDFVGEGLYQG 303
 Db 213 -RPYSFGLDFRARDNMRG---RTTVGLGLVLFQQLTGQPNVLCYA--STIFSSVGFHGG 266
 QY 304 VPRAEPTTEARRHYDEGVRMGSLGLFLOCAISLVFLSVMRLVQRCTRAVYLASVA--- 360
 Db 267 -----SSAVLASVGL---GAVKVAATLLTAMGLVDGRALLAGCALMA 308
 QY 361 -----APFVAAGATCLSHSVAVVTASAALTG---FTFSALQILPYTLASLYHR 405
 Db 309 LSVSGIGLVSPFAPVMSGSPCL--AVPNATGQGLPGDGLLODLSLPIPT-----N 360
 QY 406 EKQVFLPKYRGDTGGASSEDLSMTSLFPGPKPGAPFPNGHVAGGSGGLLPPPPPALCGASA 465
 Db 361 EDQ-----REPILSTAKTKKHPRSGDPSPAPPRALSSALPGPP----- 399
 QY 466 CDVSVRVVGEPTEARVPPRG-----ICLDLALDSAF-----LLSQVAP-- 506
 Db 400 -----LPARGHALLRWTTALLCLMVFSAPFSFGFGFTWLVLSIYPVE 442
 QY 507 -----SLFMG-----STVQLSQSVTAVMVSA--GLGLVAVY 536
 Db 443 IGRFAFCNSFWAANLFLISFLDLIGTIGLSWTFLYGLTAVLGLGYILF 496

RESULT 7

TCR1_ECOLI : STANDARD; PRT; 399 AA.
 AC P02982; :
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tetracycline resistance protein, class A (TETA(A)).
 GN TETA.
 OS Escherichia coli.
 OG Flasmid RPI.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON=Tn1721;
 RX MEDLINE=92192465; PubMed=1312499;
 RA Allmeier H., Cresnar B., Greck M., Schmitt R.;
 RT "Complete nucleotide sequence of Tn1721: gene organization and a
 RT novel gene product with features of a chemotaxis protein.";
 RL Gene 111:11-20 (1992).
 RN [2]
 RP SEQUENCE OF 85-399 FROM N.A.
 RC STRAIN=DH1;
 RA Fletcher J.N., Hart C.A., Batt R.M., Saunders J.R.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC PLASMID=RP1; TRANSPOSON=Tn1721;
 RX MEDLINE=83299270; PubMed=6310527;
 RA Waters S.H., Rogowsky P., Grinstead J., Altenbuchner J., Schmitt R.;
 RT "The tetracycline resistance determinants of RP1 and Tn1721:
 RT nucleotide sequence analysis.";
 RL Nucleic Acids Res. 11:6089-6105 (1983).
 RN [4]
 RP TOPOLOGY.
 RX MEDLINE=92388137; PubMed=1517220;
 RA Allard J.D., Bertrand K.P.;
 RT "Membrane topology of the pBR322 tetracycline resistance protein.
 RT TetA-PhoA gene fusions and implications for the mechanism of TetA
 RT membrane insertion.";

RL J. Biol. Chem. 267:17809-17819 (1992).
 CC -!- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
 CC EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
 CC ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
 CC FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
 CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X61367; CA43643.1; -
 CC EMBL; L29404; AA83545.1; -
 CC EMBL; X00006; CA24909.1; -
 CC PIR; A03509; YTECR1.
 CC PIR; JQ1479; JQ1479.
 CC InterPro; IPR004734; Drug_resist.
 CC InterPro; IPR007114; MPS.
 CC InterPro; IPR005828; Sub_transporter.
 CC InterPro; IPR005829; Sug_transporter.
 CC InterPro; IPR001958; TCR_TetA.
 CC Pfam; PF00083; sugar_tr; 1.
 CC PRINTS; PR01035; TCR_TETA.
 CC TIGRFAMs; TIGR00880; 2_A_01_02; 1.
 CC DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN 1.
 KW Antibiotic resistance; Transmembrane; Inner membrane; Transport;
 KW Antiport; Ion transport; Hydrogen ion transport; Transposable element;
 KW Plasmid.
 FT DOMAIN 1 7 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 8 27 PROBABLE.
 FT DOMAIN 28 45 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 46 66 PROBABLE.
 FT DOMAIN 67 79 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 80 100 PROBABLE.
 FT DOMAIN 101 103 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 104 124 PROBABLE.
 FT DOMAIN 125 138 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 139 159 PROBABLE.
 FT DOMAIN 160 160 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 161 181 PROBABLE.
 FT DOMAIN 182 210 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 211 231 PROBABLE.
 FT DOMAIN 232 246 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 247 267 PROBABLE.
 FT DOMAIN 268 277 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 278 298 PROBABLE.
 FT DOMAIN 299 320 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 300 320 PROBABLE.
 FT DOMAIN 321 339 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 340 360 PROBABLE.
 FT DOMAIN 361 364 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 365 385 PROBABLE.
 FT DOMAIN 386 399 CYTOPLASMIC (PROBABLE).
 FT CONFLICT 5 5 R -> I (IN REF. 3).
 FT CONFLICT 55 55 M -> V (IN REF. 3).
 FT CONFLICT 75 75 V -> I (IN REF. 3).
 FT CONFLICT 84 84 A -> T (IN REF. 3).
 FT CONFLICT 201 203 ASF -> SFV (IN REF. 3).
 SQ SEQUENCE 399 AA; 42240 MW; 298427E6B5478374 CRC64;
 Query Match 4.4%; Score 125.5; DB 1; Length 399;
 Best Local Similarity 22.3%; Pred. No. 0.051;
 Matches 104; Conservative 57; Mismatches 136; Indels 169; Gaps 27;
 QY 19 LVLNLLTFGLVCLAAAGITYVPPL---LLEVGVEKEFMT---MWLGIPVLGLVCPVLL 71
 Db 7 LIVILSTVALD---AVGIGLIMPVLPGLLRDLVHSNDVTAHVIGILLALVALMQFACAPVL 63

```
QY 72 GSASDHWGRYGR-----RPEIWSLGILLSLFLIPRAGWLAGL 112
Db 64 GALSU-----RFGRRPVLVSLAGAAVDYAIMATAPFLVLYIGRIVA----- 106
QY 113 LCPDPRPLELALLILGVGLDFGCVCFPTLEALLSDLFDPDHCRCQAYSVVAFMISLGG 172
Db 107 -----GITGATGAVA-----GAYIADITDGERARH-----GFGM-----SA 138
QY 173 CLGYLLPAIDWTSALAPYLGTQEBCLFGLLTFLITCVAATLLVAEEAALGPTPEAGL 232
Db 139 CFGFGMVA-----GPVLG-----GLM----- 154
QY 233 SAPSLPHCCPCARLAFRNLG-----ALLPRHLQCCRMPTLRFLRFLVAELCS--W--- 282
Db 155 --GGFSPH-APFFAAALNGLNFLTGCFLLPESHK-----GERPLRREALNPLASFRWARG 208
QY 283 ---MALMTFTFLFYDFVGBGLYQGVPR--EPGTEARRHYDE---GVRMGSLGLFLQCAI 334
Db 209 MTVVAALMAVFFIMQLVGQ-----VPAALWVIFGEDRFHWDATTIGISLAAGILHSLAQ 263
QY 335 SLVFSLVMDRLVQRFOTRAYSVA--APFVAAGAT--CLSHSVAVVTASAAULTGFTF 389
Db 264 AMITGPVAARLGR---RALMLGMIADGTGYILLAPATRGWMAFFIMVLLASG---GIGM 317
QY 390 SALQILPYTLASLYHREKQVFLPKYRGDTGGASSEDLSMTSFLPGP 435
Db 318 PALQAM---LSRQVBERQ-----GOLQGSIAULTSLTSIV-GP 352

RESULT 8
PUR8_STRLP STANDARD; PRT; 503 AA.
AC P42670;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Puromycin resistance protein pur8.
GN PUR8.
OS Streptomyces lipmanii (Streptomyces alboniger).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=132472;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12461;
RX MEDLINE=94109397; PubMed=7916693;
RA Terceiro J.A., Lacalle R.A., Jimenez A.;
RT "The pur8 gene from the pur cluster of Streptomyces alboniger encodes
RT a highly hydrophobic polypeptide which confers resistance to
RT puromycin.";
RL Eur. J. Biochem. 218:963-971(1993).
CC -!- FUNCTION: MAY BE INVOLVED IN ACTIVE PUROMYCIN EFFLUX ENERGIZED BY
CC A PROTON-DEPENDENT ELECTROCHEMICAL GRADIENT. IN ADDITION, IT COULD
CC BE IMPLICATED IN SECRETING N-ACETYLPUROMYCIN. THE LAST
CC INTERMEDIATE OF THE PUROMYCIN BIOSYNTHESIS PATHWAY, TO THE
CC ENVIRONMENT.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSLUCASE FAMILY).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; X76855; CAA54186.1;
DR InterPro; IPR004638; Biflux_EmrB.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
```

InterPro; IPR001411; TCR TetB.
Pfam; PF00083; sugar tx; 1.
PRINTS; PRO1036; TCRTEB.
TIGRFAMs; TIGR00711; efflux EmrB; 1.
KW Antibiotic resistance; Antibiotic biosynthesis; Transmembrane;
Transport.
FT DOMAIN 1 24 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 25 45 POTENTIAL.
FT DOMAIN 46 64 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 65 85 POTENTIAL.
FT DOMAIN 86 92 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 93 113 POTENTIAL.
FT DOMAIN 114 122 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 123 143 POTENTIAL.
FT DOMAIN 144 152 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 153 173 POTENTIAL.
FT DOMAIN 174 181 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 182 202 POTENTIAL.
FT DOMAIN 203 212 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 213 233 POTENTIAL.
FT DOMAIN 234 241 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 242 262 POTENTIAL.
FT DOMAIN 263 280 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 281 301 POTENTIAL.
FT DOMAIN 302 313 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 314 334 POTENTIAL.
FT DOMAIN 335 346 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 347 367 POTENTIAL.
FT DOMAIN 368 371 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 372 392 POTENTIAL.
FT DOMAIN 393 423 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 424 441 POTENTIAL.
FT DOMAIN 442 461 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 462 482 POTENTIAL.
FT DOMAIN 483 503 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 503 AA; 51852 MW; FA3B814DF9A3EB74 CRC64;

Query Match 4.2%; Score 121.5; DB 1; Length 503;
Best Local Similarity 23.2%; Pred No. 0.13;
Matches 131; Conservative 61; Mismatches 175; Indels 197; Gaps 32;

QY 26 FGLVCLAAAGITYV-----PALLEVGVEKFMVMV-----LGTPVLGLVCVPL 70
Db 26 WGLWILAAQLLVLDGTVNIALPSVQDLGMSDTSROWVITAYTLAFGLL-----L 79
QY 71 LGSASDHWGR-----YGRRRPFIWALSGLILLSLFLIPRAGW-LAGLL---CPDREPLEL 122
Db 80 LG-----GRVADAFGRRIE-----AVGIL-----GFLASLLGGAAPDPTLFL 119
QY 123 ALLILGVLLDFGCVCFPTLEALLSDLFDPDHCRCQAYSVVAFMISLGGCLGYLLPAID 182
Db 120 ARALQV-----FAALAPAL-ALINTLFTPEGRGKAFGVYGVGGGAUVLLAG--- 171
QY 183 WDTALAPYLGTQEBCLFGLLT--LIFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 172 -----GLLTYLDWRMCLVNPVALLALLG----- 197
QY 241 CPCRARLARLFRNLGALLPRHLQCCRMPTLRFLRFLVAEL--CSWMAIMFTFLFYTFDVG 298
Db 198 --CR-----LLPRDR-----RTGRAVRLDLPGLTLLGCGGLVAIVYAF----- 232
QY 299 GLYQGVPRAPGTEARRHYDEG--VRMGLGLFLQCAISLVFSLVMDRLVQRFOTRAYS 356
Db 233 -----AEESGMDPLVLLVGLVLMVAFALVRRVQDPLP----- 271
QY 357 ASVAAPFVAAGATCLSHSVAVVTASAAULTG-----FTFSALQILPYT--LASLVHREKQV 410
Db 272 PGVVAHRVGGG-----FLVVGLPQIGLFLFLTYLYLQGLIDYSPVLTV-----AF 320
QY 411 LPKYRGDTGGASSEDLSMTSFLPGKP-----GAPFPNGHVGGSGGLA-----PPPALC 461
Db 321 LPLGLGLIAYGSS-----LIAARLLPRTRPRTLIVGALL-----AAAGMALLTRLEPDP 370


```
Db 156 MLVSGGLALWL--ADKW-----LGWQ--MYWLMAALLIPCIATILAPE-----PT 198
Qy 227 EPAEGLSAPSLSHCCPCRLAFRLNGLALLPRLHQLCCMRPTLRRLFAVEL-----279
Db 199 D-----TIP-----VPTLEQAVVAFRPFGRN 222
Qy 280 CSWMALMTFTL-----FYTFVGEGLYQGVRAEFTGEARRHYDGVGMGLGLFLOCA 333
Db 223 NAWLILLVLYKLGDAFAMSLTTTLIRGV-----GFDAG---EVGVNKTGLLATIV 274
Qy 334 ISLVSLVNDRLVQRGTGTRAVLYASVAAPFAAGATCLSHVAVVTASAAALGFTFSALO 393
Db 275 GALTGILMQLSL-----FRALLIFI-----LQASNAGYW 307
Qy 394 ILPYTLASLYHREKQVFLPKYRGDTGCGASEDSMTSLFLPGPKGAPFNGHVAGGSGI 453
Db 308 LLSITDKHLYSMGAUVFENLCCGMGTSAFVALLMT-----LCKNFSATQPAL 356
Qy 454 LPPPPALCGASACDVSRVVVVGPEPTEARVVVGRGICLDLAILDSAFLLSQVA---PSLFW 510
Db 357 L-----SALSAVGRVVG-PVAGWVFAHGW-----STYLSVAAPVGLIL 398
Qy 511 GSIVQLSQ-----SVTAYMVS-----AAGLGLVATYFATQVVFQKSLAKYS 552
Db 399 LLVCROTLEYTRVNDNFISRTAYPAGVAFAMWTLAAGVSLAWLL-LTMDALDLTHFS 457

RESULT 10
YHJE_ECOLI
ID YHJE_ECOLI STANDARD; PRT; 440 AA.
AC P37613;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical metabolite transport protein yhjE.
GN YHJE OR B3523.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=94316500; PubMed=8041620;
RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
region from 76.0 to 81.5 minutes."
RL Nucleic Acids Res. 22:2576-2586(1994).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(potential).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. STRONG, TO
H.INFLUENZAE HI0281 AND HI0418.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U00039; AAB18499.1; -
CC EMBL; AS000429; AAC76548.1; -
CC PIR; S47743; S47743.
CC EcoGene; EG12249; yhjE.
CC InterPro; IPR004736; Cit_H_symport.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC Pfam; PF00083; sugar_1.
CC TIGRfam; TIGR00883; TRA106; 1.
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
```

```
KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
Complete proteome.
KW DOMAIN 1 34 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 35 55 1 (POTENTIAL).
FT DOMAIN 56 66 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 67 87 2 (POTENTIAL).
FT DOMAIN 88 108 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 109 129 3 (POTENTIAL).
FT TRANSMEM 130 150 4 (POTENTIAL).
FT DOMAIN 151 167 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 168 188 5 (POTENTIAL).
FT DOMAIN 189 192 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 193 213 6 (POTENTIAL).
FT DOMAIN 214 248 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 249 269 7 (POTENTIAL).
FT DOMAIN 270 289 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 290 310 8 (POTENTIAL).
FT DOMAIN 311 320 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 321 341 9 (POTENTIAL).
FT DOMAIN 342 345 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 346 366 10 (POTENTIAL).
FT DOMAIN 367 384 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 385 405 11 (POTENTIAL).
FT DOMAIN 406 410 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 411 431 12 (POTENTIAL).
FT DOMAIN 432 440 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 440 AA; 47208 MW; 3A23302A9DCB5BE CRC64;

Query Match 4.1%; Score 117; DB 1; Length 440;
Best Local Similarity 28.4%; Pred. No. 0.24;
Matches 48; Conservative 29; Mismatches 54; Indels 38; Gaps 10;

Qy 53 MTMLVIGIGVPLGVCPVLIGSADHWGRGRRRPPFIWALSIGLILSLFLI-PRAGWLAG 111
Db 295 MNAVIGFG-----VMVPVAGLLAD----AFGRKSNVIITLILFALFAPNPLG----341
Qy 112 LLCPPRRP-LELALLILGVLDLFCQVCFTPLEALLSLDFRDPHCRQAYSVAFMISL 170
Db 342 ----SGNPILVPAFLILGLSLMG----LTFGPMGALLPELF--PTEVR--YTGASPSYNV 389
Qy 171 GCGIGYLLPAIDWDTSAALPYLTQCECLFGLLTL-IFLTCVANTLLVA 218
Db 390 ASILG-----ASVAPYIAAWLQNTNYGILGAVGLYLAANAGLTLLA 428

RESULT 11
SMVA_SALTY
ID SMVA_SALTY STANDARD; PRT; 495 AA.
AC P37594;
DT 01-OCT-1994 (Rel. 30, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Methyl viologen resistance protein smva.
GN SMVA OR STM1574.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SL1303;
RX MEDLINE=95011654; PubMed=7926834;
RA Hongo E., Morimyo M., Mita K., Machida I., Hama-Inaba H., Tsuji H.,
RA Ichimura S., Noda Y.;
RT "The methyl viologen-resistance-encoding gene smva of Salmonella
typhimurium."
RL Gene 148:173-174(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.B., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
```


Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;
 "Complete genome sequence of *Salmonella enterica* serovar Typhimurium Lf2.", 413:852-856(2001).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Potential).
 CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D26057; BAA05055.1; -
 CC EMBL; AE008769; AAL20492.1; -
 CC StryGene; SG10384; smva
 CC InterPro; IPR007114; MFS.
 CC InterPro; IPR001411; TCR-TetB.
 CC PRINTS; PR01036; TCR-TetB.
 KW Transmembrane; Inner membrane; Complete proteome.
 FT TRANSMEM 5 25 POTENTIAL.
 FT TRANSMEM 44 64 POTENTIAL.
 FT TRANSMEM 73 93 POTENTIAL.
 FT TRANSMEM 96 116 POTENTIAL.
 FT TRANSMEM 135 155 POTENTIAL.
 FT TRANSMEM 158 178 POTENTIAL.
 FT TRANSMEM 192 212 POTENTIAL.
 FT TRANSMEM 220 240 POTENTIAL.
 FT TRANSMEM 260 280 POTENTIAL.
 FT TRANSMEM 299 319 POTENTIAL.
 FT TRANSMEM 327 347 POTENTIAL.
 FT TRANSMEM 357 377 POTENTIAL.
 FT TRANSMEM 391 411 POTENTIAL.
 FT TRANSMEM 469 489 POTENTIAL.
 FT CONFLICT 144 144 A -> R (IN REF. 1).
 FT CONFLICT 147 147 G -> V (IN REF. 1).
 FT CONFLICT 182 182 V -> D (IN REF. 1).
 FT CONFLICT 198 198 A -> V (IN REF. 1).
 FT CONFLICT 227 227 F -> Y (IN REF. 1).
 FT CONFLICT 447 447 L -> Y (IN REF. 1).
 FT CONFLICT 454 457 AILD -> GYLT (IN REF. 1).
 FT SEQUENCE 495 AA; 52134 MW; 961F7C748CDA164 CRC64;
 Query Match 4.1%; Score 117; DB 1; Length 495;
 Best Local Similarity 21.0%; Pred. No. 0.27;
 Matches 92; Conservative 60; Mismatches 151; Indels 136; Gaps 22;
 QY 17 QLLVNLVLFGL-----EVCLAAGITVYVPLLLEVG-----EKFMT 54
 DB 72 RLILMGTLFGLASLAASHTASWLIATRVLLAIGAAIMVPAVL-AGIRATPCSEKHN 130
 QY 55 MVLGI-----GPVLG-----IVCVPL-----IGSADHWGRYGR- 85
 DB 131 MALGVAAVVGSGAAGFGLTGILLBHFYWGVSFLNVLVVMGLTARYVPRQAGRD 190
 QY 86 RPP-----IWALS-----GILLSLFLIPRAGWLAGL 112
 DB 191 QPLNLGHAVNLIITAILLVYSAKTALKHLSLWISFTLLTGALLGLFIRQLATSRPM 250
 QY 113 LCPDPRPLEALILVGLLDPCGVCFTPLEALLSDLFRDPDCHQRAYSVAFMI----- 168
 DB 251 I--DMRLFTHRIILSGV-VNMAWTAMITLVGFELLMQAQLQV-HGLSPYEAGVFMVPMV 306
 QY 169 -----SLGGCL-----GYLLPAIDWDTSAAPYLGTOECLFGLLTILFLTC 210
 DB 307 ASFGSGFIAGVLVSRILGLRLVATGGMALSAISPYGLAMTDF-STQOQWAGMLMALIGFS- 364

QY 211 VVATLLVBEAALGPTPEAE-GLSAPSLSPHCCPCRARLAFRLNLGALLRLHOLCCMRP 269
 DB 365 -AASALLASTSAIMAAAPAKAAAGAIETMAYELGAGLGIAIFGLLSRFSASIRLPA 423
 QY 270 TLRLFLVACLSCWMLMTTFLFYTFVGG--LYQGV--RAEPGTGEARRH---YDEGVR 322
 DB 424 GLEAQIARASSM-----GEAVQLANSLPPTQQAILDAARHAFINSHVA 470
 QY 323 MGSLG-LFLOCAISLVFSL 340
 DB 471 LSSAGSMILLAVGMWFSL 489
 RESULT 12
 Y051 MYCTU
 ID Y051 MYCTU STANDARD; PRT; 640 AA.
 AC Q10880; O53628;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein Rv0083.
 GN Rv0083 OR MT0090 OR MTCY251.01 OR MTV030.27.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekai F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares R., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE OF CHLOROPLASTS OR MITOCHONDRIA.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL021428; CAAL6264.1; -
 CC EMBL; AE006920; AAK44315.1; -
 CC PIR; D70850; D70850.
 CC TIGR; MT0090; -
 CC TubercuList; Rv0083; -
 CC InterPro; IPR003918; NADH_oxred.
 CC InterPro; IPR001750; Oxidored_q1.
 CC Pfam; PF00361; oxidored_q1; 1.
 CC PRINTS; PR01437; NUOXDRDTASE4.

KW Hypothetical protein; Oxidoreductase; Transmembrane;
 KW Complete proteome.
 FT TRANSMEM 8 28 POTENTIAL.
 FT TRANSMEM 52 72 POTENTIAL.
 FT TRANSMEM 90 110 POTENTIAL.
 FT TRANSMEM 136 156 POTENTIAL.
 FT TRANSMEM 179 199 POTENTIAL.
 FT TRANSMEM 208 228 POTENTIAL.
 FT TRANSMEM 241 261 POTENTIAL.
 FT TRANSMEM 277 297 POTENTIAL.
 FT TRANSMEM 298 318 POTENTIAL.
 FT TRANSMEM 352 372 POTENTIAL.
 FT TRANSMEM 391 411 POTENTIAL.
 FT TRANSMEM 446 466 POTENTIAL.
 FT TRANSMEM 497 517 POTENTIAL.
 FT TRANSMEM 619 639 POTENTIAL.
 SQ SEQUENCE 640 AA; 65627 MW; 6254C289DBD108A8 CRC64;
 Query Match 3.9%; Score 111.5; DB 1; Length 640;
 Best Local Similarity 22.3%; Pred. No. 0.87;
 Matches 129; Conservative 73; Mismatches 222; Indels 155; Gaps 29;
 QY 16 AQLLV---NLTFLG--EVCLAAGTIVY---PPLLEVGVEKFTMTVLGIGPVLG 64
 Db AAMLVPAAGSVTTFLAWELMAISLILVLSHARPQVRSAGLWYAVMTQLGFIALVVG 152
 QY 65 LVCVPLGASDHWGRYGRRRPFIWALSIGLILSLF-----LIPRAGHLAGL 112
 Db LVLVLAAG--GSDRFAG-----LGAVCDGVRAAVFMTLVGFGSKAGLVLPHAWLPRA 203
 QY 113 LCPDPRP---LELALLILGV-GLLDFCGQVCTPLEALLSDLFDPDHCROAYSVYAFM 167
 Db HPEAPSPVSAALMAAMVNGIYGVFDLQ-----LCPGRWGLA-----L 246
 QY 168 ISLGG---CLGYLLPAIDMTSALAPYLGTQBECLFGLLTLPITCVAATLVAEBAALG 224
 Db LAVGTSALYGVLAQSVAAADLKELAYSTTEN---MGLITL-----ALGAATLFDATGAYG 299
 QY 225 PTEPAGLSAPSUSPHCCPCRARLAPRNLGALLPR-----LHOLCCRMPTLRLFLV 276
 Db PASIAAAAAAMLHMAHA--AFKSLAFMAAGSVLAATGLRDLGLGLARRMPAT----- 351
 QY 277 AELCSMWALMTFLFY-----TDFVGEV-LYGVVPRAAPGTEARRHYDEGVR 322
 Db -----TVFVGVAALGACGLPLGAGFVSEMLVQSLIHAAPGHPDPIVALTTPLA 399
 QY 323 MSLGLFLQCAISLVFS-----LVMDRLVQRFGRTRAVYLASVAAPFVAAGATCLSHSVA 376
 Db VGVVALATGLSVAAMTKATFGIGFLARPRSTQAEAREAPASMEAGMAIAAGA-CLVLAVA 458
 QY 377 --VVTASALTGTFPSALQILPVT-LASLYHREKQVFLPKYRGDTGGASDSLMTSFLP 433
 Db PLLVAPMVRRAAATLPAQAQVKTGLGAV-----VRLPAMSG-----SIAP 499
 QY 434 GPKPGAPFPNGHVGAGSGLLP-----PPPA-----LCGASACDVSVRVVVGEPTEAR 481
 Db GVTAIAAVL-----AAALAVAVLARWRFRPPAPARLPFLWACGAA--DLTVRMQYATSPAE 553
 QY 482 VFPGRCICLDLALDSAFLLSQVAPSLFMGSIQVLSQSV 520
 Db --FLQRFVGDVLRPDITDVTHTAESRYMAERTYRTAV 590
 RESULT 13
 PHDK NOCSK
 ID PHDK NOCSK STANDARD; PRT; 473 AA.
 AC O24723;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable 1-hydroxy-2-naphthoate transporter.
 OS Nocardioides sp. (strain KP7).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Propionibacterineae; Nocardioidaceae; Nocardioides.
 NCBI_TaxID=35761;
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=97474276; PubMed=9335300;
 RA Iwabuchi T., Harayama S.;
 RT "Biochemical and genetic characterization of 2-carboxybenzaldehyde
 dehydrogenase, an enzyme involved in phenanthrene degradation by
 Nocardioides sp. strain KP7";
 RL J. Bacteriol. 179:6488-6494 (1997).
 CC -!- FUNCTION: PROBABLY INVOLVED IN THE TRANSPORT OF 1-HYDROXY-2-
 NAPHTHOATE.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 (Potential).
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to a license@isb-sib.ch).
 CC
 CC -----BAA23264.1; --
 DR EMBL; AB000735; BAA23264.1; --
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
 KW Transport; Transmembrane; Inner membrane.
 FT DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 26 46
 FT DOMAIN 47 59 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 60 80 2 (POTENTIAL).
 FT DOMAIN 81 91 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 92 112 3 (POTENTIAL).
 FT DOMAIN 113 121 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 122 142 4 (POTENTIAL).
 FT DOMAIN 143 153 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 154 174 5 (POTENTIAL).
 FT DOMAIN 175 180 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 181 201 6 (POTENTIAL).
 FT DOMAIN 202 233 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 234 284 7 (POTENTIAL).
 FT DOMAIN 285 304 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 305 325 8 (POTENTIAL).
 FT DOMAIN 326 330 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 331 351 9 (POTENTIAL).
 FT DOMAIN 352 354 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 355 375 10 (POTENTIAL).
 FT DOMAIN 376 398 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 399 418 11 (POTENTIAL).
 FT DOMAIN 419 421 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 422 444 12 (POTENTIAL).
 FT DOMAIN 445 473 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 473 AA; 49109 MW; D6D765D376260D8A CRC64;
 Query Match 3.9%; Score 111; DB 1; Length 473;
 Best Local Similarity 24.3%; Pred. No. 0.69;
 Matches 114; Conservative 66; Mismatches 180; Indels 110; Gaps 27;
 QY 12 RHRKAQLLVNLTFF--GLEVCLAGITVPPILLLEVGVEEKFTMTVLGIGPVGLVCVP 69
 Db RQRIIVLMVFLIMVADGMDITLASHL--FPVIRDWGVFVSATVLVSLGVVMAIGAL 79
 QY 70 LLGSASDHWGRYGRRRPFIWALSIGLILSLFLIPRAG-WLAGLLCPDRPLELALLIG 128
 Db VSGPVDNRW-GRKG-----VTVGVFV--LFLATAGLGLTG----DIHSPALRIISC 125
 QY 129 VGLLDPCGQVCTPLEALLSDLFDPDHCRC-QAYSVYAFMISLGGCLGLYLLPAIDWDTSA 187

Db 126 FGL-----GAVMPVAL-TIVADMM--PKARRAQMVSIATFAGVGVGSIIGAYLAA-----A 172
Qy 188 LAPVLGTQEBCLF-GLTLTFLTCVAATLLVABEA-----ALGPTBPAG 231
Db 173 VIPTLQWQVVLATLAPLILIPFFVA--LVPEPAIISVVRGIPPEARIRSAALALVAPDRD 230
Qy 232 LSAPSLSPHCCPCRARAFNIGALLPRLHOLCCMRPTRLRRLFVABLCSWMLAMTFTTLF 291
Db 231 IAGVDLT-----RAGL-----TLGAGEVR-----AKALFAELICRPLLGVTLLIW 270
Qy 292 YTFPVGEG-----LYQVGP-----RAEPCTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMD 343
Db 271 GVFFVQSGLLVLQVMPMLQAPAGLST---VESGLIVAMTGWG-----ALIGQUTTA 332
Qy 344 RLQVRFQFTRAVYLASVAAFPVAAGATCLSHSVAVVTASAAALTGTF-----SALQI-L 395
Db 323 FILKRF-DRIALAAFIWVS-----VGLLIVAAFGTGFGRFGYFTLFAIGLSL 371
Qy 396 PYTLASLYHREKQVFLPKYR-----GDTGGASSDSLSMTSLFPGKPGAPF 441
Db 372 PATAAAMQSVTTLAYEEPRATGWSAGFAGRLGTLTYGALGOTLIGAGF 421

RESULT 14
CAIA_HUMAN STANDARD; PRT; 680 AA.
AC Q03692;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Collagen alpha 1(X) chain precursor.
GN COL10A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=92109659; PubMed=1764025;
RA Thomas J.T., Stesswell C.J., Rash B., Nicolai H., Jones T.,
RA Solomon E., Grant M.E., Boot-Handford R.P.;
RT "The human collagen X gene. Complete primary translated sequence and
RT chromosomal localization."
RL Biochem. J. 280:617-623(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93012005; PubMed=1397333;
RA Reichenberger E., Beier F., Luvalle P., Olsen B.R., von der Mark K.,
RA Bertling W.M.;
RT "Genomic organization and full-length cDNA sequence of human collagen
RT X."
RL FEBS Lett. 311:305-310(1992).
RN [3]
RP SEQUENCE FROM N.A.
RA Beier F., Lammi M.B., von der Mark K.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Williams S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 52-680 FROM N.A.
RX MEDLINE=92267014; PubMed=1587271;
RA Apte S.S., Seldin M.F., Hayashi M., Olsen B.R.;
RT "Cloning of the human and mouse type X collagen genes and mapping of
RT the mouse type X collagen gene to chromosome 10."
RL Eur. J. Biochem. 206:217-224(1992).
RN [6]
RP SEQUENCE OF 561-666 FROM N.A.
RX MEDLINE=91243838; PubMed=2037056;
RA Apte S., Mattei M.-G., Olsen B.R.;
RT "Cloning of human alpha 1(X) collagen DNA and localization of the
RT COL10A1 gene to the q21-q22 region of human chromosome 6.";

RL FEBS Lett. 282:393-396(1991).
RN [7]
RP SEQUENCE OF 547-655 FROM N.A.
RX MEDLINE=92077285; PubMed=1743401;
RA Reichenberger E., Aigner T., von der Mark K., Stoeb H., Bertling W.;
RT "In situ hybridization studies on the expression of type X collagen
RT in fetal human cartilage."
RL Dev. Biol. 148:562-572(1991).
RN [8]
RP REVIEW ON VARIANTS.
RX MEDLINE=97255959; PubMed=9101290;
RA Kuivaniemi H., Tromp G., Prockop D.J.;
RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
RT associated collagen (type IX), and network-forming collagen (type X)
RT cause a spectrum of diseases of bone, cartilage, and blood vessels."
RL Hum. Mutat. 9:300-315(1997).
RN [9]
RP VARIANTS SMCD ASP-598 AND PRO-614.
RX MEDLINE=94136476; PubMed=8304336;
RA Wallis G.A., Rash B., Sweetman W.A., Thomas J.T., Super M., Evans G.,
RA Grant M.E., Boot-Handford R.P.;
RT "Amino acid substitutions of conserved residues in the
RT carboxyl-terminal domain of the alpha 1(X) chain of type X collagen
RT occur in two unrelated families with metaphyseal chondrodysplasia
RT type Schmid."
RL Am. J. Hum. Genet. 54:169-178(1994).
RN [10]
RP VARIANT SMCD ARG-591.
RX MEDLINE=94272470; PubMed=8004099;
RA McIntosh I., Abbott M.H., Warman M.L., Olsen B.R., Francomano C.A.;
RT "Additional mutations of type X collagen confirm COL10A1 as the
RT Schmid metaphyseal chondrodysplasia locus."
RL Hum. Mol. Genet. 3:303-307(1994).
RN [11]
RP VARIANT SMCD VAL-618.
RX MEDLINE=95181449; PubMed=7876225;
RA Chan D., Cole W.G., Rogers J.G., Bateman J.F.;
RT "Type X collagen multimer assembly in vitro is prevented by a Gly618
RT to Val mutation in the alpha 1(X) NC1 domain resulting in Schmid
RT metaphyseal chondrodysplasia."
RL J. Biol. Chem. 270:4558-4562(1995).
RN [12]
RP VARIANTS SMCD ARG-545; GLU-595; HIS-597; LYS-617; ARG-644 AND GLY-648.
RX MEDLINE=95331767; PubMed=7607655;
RA Bonaventure J., Chaminade F., Maroteaux P.;
RT "Mutations in three subdomains of the carboxy-terminal region of
RT collagen type X account for most of the Schmid metaphyseal
RT dysplasias."
RL Hum. Genet. 96:58-64(1995).
RN [13]
RP VARIANT SMCD PRO-600.
RX MEDLINE=96375754; PubMed=8782043;
RA Wallis G.A., Rash B., Sykes B., Bonaventure J., Maroteaux P.,
RA Zabel B., Wynne-Davies R., Grant M.E., Boot-Handford R.P.;
RT "Mutations within the gene encoding the alpha 1(X) chain of type X
RT collagen (COL10A1) cause metaphyseal chondrodysplasia type Schmid but
RT not several other forms of metaphyseal chondrodysplasia."
RL J. Med. Genet. 33:450-457(1996).
RN [14]
RP VARIANTS SMCD GLU-18 AND ARG-18.
RX MEDLINE=97220591; PubMed=9067753;
RA Ikegawa S., Nakamura K., Nagano A., Haga N., Nakamura Y.;
RT "Mutations in the N-terminal globular domain of the type X collagen
RT gene (COL10A1) in patients with Schmid metaphyseal
RT chondrodysplasia."
RL Hum. Mutat. 9:131-135(1997).
RN [15]
RP VARIANTS SMD GLU-595.
RX MEDLINE=99057503; PubMed=9837818;
RA Ikegawa S., Nishimura G., Nagai T., Hasegawa T., Ohashi H.,
RA Nakamura Y.;
RT "Mutation of the type X collagen gene 'COL10A1' causes
RT spondylometaphyseal dysplasia.";

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 17:16:46 ; Search time 40 Seconds
(without alignments)
3567.577 Million cell updates/sec

Title: US-09-593-793a-113
Perfect score: 2861
Sequence: 1 MVQRLWVSRLLRHRKAQLL.....AIYFATQVVPDKSLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2861	100.0	553	4	Q96JT2
2	2798	97.8	553	6	Q95KI5
3	2602	90.9	553	11	Q8XOH7
4	2531	88.5	501	6	Q95KC5
5	2099	73.4	450	11	Q8K252
6	519.5	18.2	599	5	Q9VSV1
7	512	17.9	112	11	Q8RI10
8	484.5	16.9	530	11	Q8C204
9	456	15.9	576	13	Q90Z74
10	446	15.6	751	11	Q8K4S3
11	394.5	13.8	754	5	Q8MMT9
12	347.5	12.1	515	10	Q65803
13	347.5	12.1	515	10	Q9FNR6
14	346.5	12.1	501	10	Q9SQK6
15	346.5	12.1	515	10	Q9SLN7
16	346	12.1	612	10	Q9SQK5

ALIGNMENTS

RESULT 1

Q96JT2 ID Q96JT2 PRELIMINARY; PRT; 553 AA.

AC Q96JT2; DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE Prosteln.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Prostate;

RX MEDLINE=21139094; PubMed=11245466;

RA Xu J., Kalos M., Stolk J.A., Zasloff E.J., Zhang X., Houghton R.L.,

RA Filho A.M., Nolasco M., Badaro R., Reed S.G.;

RT "Identification and characterization of prostein, a novel prostate-

RT specific protein."

RL Cancer Res. 61:1563-1568(2001).

DR EMBL: AV033593; AAK54386.1; --

SQ SEQUENCE 553 AA; 59322 MW; 0AFA23FBC742A667 CRC64;

Query Match 100.0%; Score 2861; DB 4; Length 553;

Best Local Similarity 100.0%; Pred. No. 2.6e-209;

Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLNVLLTFGLVCLAAAGITVPPLLLLEVGVVEEKFTWVLGIG 60

DB 1 MVQRLWVSRLLRHRKAQLLNVLLTFGLVCLAAAGITVPPLLLLEVGVVEEKFTWVLGIG 60

QY 61 PVLGLVCPPLGASDHWGRYGRRRPTIWSLGLLSLFLIPRAGWLAGLLCPDPRPL 120

DB 61 PVLGLVCPPLGASDHWGRYGRRRPTIWSLGLLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLILGVLLDFCGQVCFTPLEALLSDLFRDPDHCQAYSVYAFMISLGCGLYLLPA 180

DB 121 ELALLILGVLLDFCGQVCFTPLEALLSDLFRDPDHCQAYSVYAFMISLGCGLYLLPA 180

Q9S725 apium grave
O04077 vicia faba
Q9SP63 vitis vinif
Q92TB9 apium grave
Q8VYX3 brassica ol
Q43653 solanum tub
Q9XHL6 pisum sativ
Q944W2 oryza sativ
O80605 arabidopsis
Q948L0 oryza sativ
Q9FV16 lycopersico
Q9SP14 alonsoa mer
Q8RWG6 arabidopsis
Q39231 arabidopsis
O80550 arabidopsis
Q40583 nicotiana t
Q9FVJ6 lycopersico
Q9EXM0 zea mays (m
Q41152 ricinus com
Q81PM4 citrus sine
Q9FV92 solanum tub
Q8VYX4 brassica ol
Q81PM6 citrus sine
Q9M422 hordeum vul
Q65929 daucus caro
Q9SP15 asarina bar
Q8RUL3 triticum ae
Q40167 lycopersico

QY 181 IDWTSALAPYLGTQBECLFGLLTLLFTCTVAATLLVAEEAALGPTEPAEGLSAPLSLSPH 240
 DB 181 IDWTSALAPYLGTQBECLFGLLTLLFTCTVAATLLVAEEAALGPTEPAEGLSAPLSLSPH 240
 QY 241 CCPCARLARAFNLGALLPRLHQLCCMPRTLRLRFLVAELCSWMLMTFTFLFYTDFVGEGL 300
 DB 241 CCPCARLARAFNLGALLPRLHQLCCMPRTLRLRFLVAELCSWMLMTFTFLFYTDFVGEGL 300
 QY 301 YQGVPAEPTGTEARRHYDEGVMSGLFLQCAISLVFSLVMDRLVQRTGTRAVYLASVA 360
 DB 301 YQGVPAEPTGTEARRHYDEGVMSGLFLQCAISLVFSLVMDRLVQRTGTRAVYLASVA 360
 QY 361 APVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 APVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
 DB 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
 QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMWSAAGLGLVAYIFATQ 540
 DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMWSAAGLGLVAYIFATQ 540
 QY 541 VVFDKSDLAKYSA 553
 DB 541 VVFDKSDLAKYSA 553

RESULT 2

Q95K15 PRELIMINARY; PRT; 553 AA.
 ID Q95K15
 AC Q95K15
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 59.4 kDa protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Temporal cortex;
 RA Oeada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB060851; BAB46871.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 553 AA; 59392 MW; 0718F3A91FB3BF1E CRC64;

Query Match 97.8%; Score 2798; DB 6; Length 553;
 Best Local Similarity 98.2%; Pred. No. 1.6e-204;
 Matches 542; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MVQRLVWSRLRRKRAQLLLVLLTFGLVCLAAAGITYVPPILLVGVBEKFMVVLGIG 60
 DB 1 MVQRLVWSRLRRKRAQLLLVLLTFGLVCLAAAGITYVPPILLVGVBEKFMVVLGIG 60
 QY 61 PVGLVVCVPLGASDHWGRYGRRRPFTWALSGLLSLFLIPRAGWLAGLCPDPRPL 120
 DB 61 PVGLVVCVPLGASDHWGRYGRRRPFTWALSGLLSLFLIPRAGWLAGLCPDPRPL 120
 QY 121 ELALLTILGVCLDPCQVCFPLEALLSDLFRDPDHCRAQYSYAFWISLGGCLGYLLPA 180
 DB 121 ELALLTILGVCLDPCQVCFPLEALLSDLFRDPDHCRAQYSYAFWISLGGCLGYLLPA 180
 QY 181 IDWTSALAPYLGTQBECLFGLLTLLFTCTVAATLLVAEEAALGPTEPAEGLSAPLSLSPH 240

DB 181 IDWTSALAPYLGTQBECLFGLLTLLFTCTVAATLLVAEEAALGPTEPAEGLSAPLSLSPH 240
 QY 241 CCPCARLARAFNLGALLPRLHQLCCMPRTLRLRFLVAELCSWMLMTFTFLFYTDFVGEGL 300
 DB 241 CCPCARLARAFNLGALLPRLHQLCCMPRTLRLRFLVAELCSWMLMTFTFLFYTDFVGEGL 300
 QY 301 YQGVPAEPTGTEARRHYDEGVMSGLFLQCAISLVFSLVMDRLVQRTGTRAVYLASVA 360
 DB 301 YQGVPAEPTGTEARRHYDEGVMSGLFLQCAISLVFSLVMDRLVQRTGTRAVYLASVA 360
 QY 361 APVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 APVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
 DB 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
 QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMWSAAGLGLVAYIFATQ 540
 DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMWSAAGLGLVAYIFATQ 540
 QY 541 VVFDKSDLAKYSA 553
 DB 541 VVFDKSDLAKYSA 553

RESULT 3

Q8K0H7 PRELIMINARY; PRT; 553 AA.
 ID Q8K0H7
 AC Q8K0H7
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein (PROSTEIN homolog).
 GN 2210413P12RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 DR EMBL; BC031381; AAH31381.1; -;
 DR EMBL; AK035428; BAC29063.1; -;
 DR MGD; MGI:1922082; 2210413P12RIK.
 DR InterPro; IPR005805; Rieseke.
 DR PROSITE; PS00200; RIESEK_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 553 AA; 59742 MW; 6A474PEA630B4B13 CRC64;

Query Match 90.9%; Score 2602; DB 11; Length 553;
 Best Local Similarity 90.9%; Pred. No. 1.3e-189;
 Matches 502; Conservative 14; Mismatches 36; Indels 0; Gaps 0;

QY 1 MVQRLVWSRLRRKRAQLLLVLLTFGLVCLAAAGITYVPPILLVGVBEKFMVVLGIG 60
 DB 1 MVQRLVWSRLRRKRAQLLLVLLTFGLVCLAAAGITYVPPILLVGVBEKFMVVLGIG 60
 QY 61 PVGLVVCVPLGASDHWGRYGRRRPFTWALSGLLSLFLIPRAGWLAGLCPDPRPL 120
 DB 61 PVGLVVCVPLGASDHWGRYGRRRPFTWALSGLLSLFLIPRAGWLAGLCPDPRPL 120


```

121 ELALLILGVGLDFCGVCFPTLEALLSDFRDPDCHCRQAVSVYAFMISLGGCLGYLLPA 180
121 ELALLILGVGLDFCGVCFPTLEALLSDFRDPDCHCRQAVSVYAFMISLGGCLGYLLPA 180
181 IDWDTMALAPYLGTOECLFGLTLIFLTCVAATLLVAEEAALGTPPAEGLSPH 240
181 IDWDTMALAPYLGTOECLFGLTLIFLTCVAATLLVAEEAALGTPPAEGLSPH 240
241 CCPCRAFLAFNLCALLPRHLQCCMRPTRLRLRFLVAELCSWMAALMTFTFYDFVGEGL 300
241 CCPCRAFLAFNLCALLPRHLQCCMRPTRLRLRFLVAELCSWMAALMTFTFYDFVGEGL 300
301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQFGTRVYLASVA 360
301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQFGTRVYLASVA 360
361 APVAAAGATCLSHSVAVVTASAAITGTFSSALQILPYTLASLYHREKQVFLPKYRGDTG 420
361 TFPVAAAATCLSHSVAVVTASAAITGTFSSALQILPYTLASLYHREKQVFLPKYRGDTG 420
421 ASSDLSMTSFLPGPKGAPPNGHVGAGGSLGPPPPALCGASACDVSVRVVGEPTA 480
421 SSGEDSOTTFPLPGPKGAPPNGHVGAGGSLGPPPPALCGASACDVSVRVVGEPTA 480
481 RVVPGRGICLDLALDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYPATQ 540
481 RVVPGRGICLDLALDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYPATQ 540
541 VVFDKSLDIAKYS 552
541 VVFDKSLDIAKYS 552

RESULT 4
Q95KCS PRELIMINARY; PRT; 501 AA.
AC Q95KCS;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE Hypothetical 53.4 kDa protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Medulla oblongata;
RA Osada N., Hida M., Kuwada J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062977; BAB60745.1;
KW Hypothetical protein.
SQ SEQUENCE 501 AA; 53447 MW; 8C554BBD04EE0470 CRC64;

Query Match 88.5%; Score 2531; DB 6; Length 501;
Best Local Similarity 97.8%; Pred. No. 2.8e-184;
Matches 489; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 53 MTWVLGIPVLGVPLVGLSGASDHWGRGRRRPTWALSGLILSLFLIPRAGWLAGL 112
DB 1 MTWVLGIPVLGVPLVGLSVPLVGLSGASDHWGRGRRRPTWALSGLILSLFLIPRAGWLAGL 60
QY 113 LCPDPRPLEALLILGVGLDFCGVCFPTLEALLSDFRDPDCHCRQAVSVYAFMISLGG 172
DB 61 LCPDPRPLEALLILGVGLDFCGVCFPTLEALLSDFRDPDCHCRQAVSVYAFMISLGG 120
QY 173 CLGYLLPAIDWDTMALAPYLGTOECLFGLTLIFLTCVAATLLVAEEAALGTPPAEGL 232

```

```

121 CLGYLLPAIDWDTMALAPYLGTOECLFGLTLIFLTCVAATLLVAEEAALGTPPAEGL 180
233 SPSLSHCHCCPCRAFLAFNLCALLPRHLQCCMRPTRLRLRFLVAELCSWMAALMTFTFY 292
181 SPSLSHCHCCPCRAFLAFNLCALLPRHLQCCMRPTRLRLRFLVAELCSWMAALMTFTFY 240
293 TDFVGBGLYQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQFGTR 352
241 TDFVGBGLYQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQFGTR 300
353 AVYLASVAAPVAAAGATCLSHSVAVVTASAAITGTFSSALQILPYTLASLYHREKQVFLP 412
301 AVYLASVAAPVAAAGATCLSHSVAVVTASAAITGTFSSALQILPYTLASLYHREKQVFLP 360
413 KYRGDTGGASSEDLSMTSFLPGPKGAPPNGHVGAGGSLGPPPPALCGASACDVSVRV 472
361 KYRGDTGGASSEDLSMTSFLPGPKGAPPNGHVGAGGSLGPPPPALCGASACDVSVRV 420
473 VUGEPTEARVVPGRGICLDLALDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGL 532
421 VUGEPTEARVVPGRGICLDLALDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGL 480
533 VAIYPATQVVDKSLDIAKYS 552
481 VAIYPATQVVDKSLDIAKYS 500

RESULT 5
Q8K252 PRELIMINARY; PRT; 450 AA.
ID Q8K252;
AC Q8K252;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
GN 2210413P12RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC034084; AAH34084.1;
DR MGI; MGI:1922082; 2210413P12RIK.
DR InterPro; IPR005805; Rieseke.
DR PROSITE; PS00200; RIESEKE_2; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 450 AA; 48261 MW; 72C703AEC131302C CRC64;

Query Match 73.4%; Score 2099; DB 11; Length 450;
Best Local Similarity 90.0%; Pred. No. 1.7e-151;
Matches 404; Conservative 12; Mismatches 33; Indels 0; Gaps 0;

QY 104 PRAGWLAGLCPDPRPLEALLILGVGLDFCGVCFPTLEALLSDFRDPDCHCRQAVSV 163
DB 1 PRAGWLAGLCPDPRPLEALLILGVGLDFCGVCFPTLEALLSDFRDPDCHCRQAVSV 60
QY 164 YAFMISLGGCLGYLLPAIDWDTMALAPYLGTOECLFGLTLIFLTCVAATLLVAEEAAL 223
DB 61 YAFMISLGGCLGYLLPAIDWDTMALAPYLGTOECLFGLTLIFLTCVAATLLVAEEAAL 120
QY 224 GTEPAEGLSAPSLSPHCCPCRAFLAFNLCALLPRHLQCCMRPTRLRLRFLVAELCSWMA 283
DB 121 GTEPAEGLSAPSLSPHCCPCRAFLAFNLCALLPRHLQCCMRPTRLRLRFLVAELCSWMA 180
284 ALMTFTFYDFVGBGLYQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMD 343
DB 181 ALMTFTFYDFVGBGLYQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMD 240
QY 344 RLVRFGTRAVYLASVAAPVAAAGATCLSHSVAVVTASAAITGTFSSALQILPYTLASLY 403

```

```

Db 241 RLKQKGTSLVSLASVMTFPVAAAATCLSHVVVVTASALTGTFTSALQILYPTLASLY 300
QY 404 HREKQVFLPKYRGDTGASSEDLSMTSFLPGPKGAPFPNGHVGAGSGLLPPPPALCGA 463
Db 301 HREKQVFLPKYRGDAGSGSGEDSQTTSFLPGPKGALFPNGHVGSGSGILAPPPALCGA 360
QY 464 SACDVSUVVVGPTARVPGRGICLDLAILDSAFLLSVAFLSGMSVQLSOSVTAY 523
Db 361 SACDVSUVVVGPTARVPGRGICLDLAILDSAFLLSVAFLSGMSVQLSHSVTAY 420
QY 524 MVSAAAGLGLVAIFYATQVVPDKSLAKYS 552
Db 421 MVSAAAGLGLVAIFYATQVVPDKSLAKYS 449

RESULT 6
QVSV1 PRELIMINARY; PRT; 599 AA.
AC QVSV1
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CG4484 protein.
GN CG4484.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKLEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan M.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003552; AAF50310.1; -

```

```

DR FlyBase; FBgn0035968; CG4484.
SQ SEQUENCE 599 AA; 66057 MW; C5381D334CFBF2EB CRC64;

Query Match 18.2%; Score 519.5; DB 5; Length 599;
Best Local Similarity 23.7%; Pred. No. 2,7e-31;
Matches 150; Conservative 100; Mismatches 201; Indels 183; Gaps 14;

QY 12 RHKQALLAVNLTTFGLVCLAAGITVVPPLLEVGVEEKFMTVLGIGVGLGVCVPLL 71
Db 46 RTRPFMFSLSAIAMAIEFAYAEYSFVSPILLQIGVDHKKMSMTWGLSFLIGFMSPLL 105
QY 72 GSASDHWRGRRRRPFIWALSIGILLSLFLIP-----R 105
Db 106 GSISDRCKLRWGRRRPIIILSFGIMCGLILVPGKDLGILLGADAGTYAESALNTSSS 165
QY 106 AGWLAGLLC-----PDRPLELA--LLILGVGLDFCGQVCPFPLEALLSDFLRDPDHC 157
Db 166 GGSVAALVSGEATGSPASDYKFAVILITLGMVLLDFDADTCQTPARTYLLDMC--VPEEQ 224
QY 158 ROAYSVYAFMISLGCLGYLLPAIDWDTSAALPYLGTQECFLGLTLFLTCTVAATL-- 215
Db 225 PKAMTMFAALFAGFGGTGTYAIGVDWEITHIGSFMGGNIPVFTLVTIILFAVCYLITVT 284
QY 216 -----LVAEEAALGP-----TBPABG 231
Db 285 FREIPLPLIEQDELRLPLSEQAIKKELKKNTIYVIOETQLELQMSDDPKRLAALQG 344
QY 232 LSAPLSLP-----HCCPCRLARLAFNLGALLPRLHQLCCRMPTLRLLF 275
Db 345 SYONGYSPAVEKQKGSQDLTQSDYDAPVSLKAYLKSIFI-----NPYSMRMLA 393
QY 276 VAEELCSWALMTFTLYTDFVGEGLYQVPRAPGTEARRHYDEGVYRMSGLGLFLQCAIS 335
Db 394 LTNLFCEWGMGHVITYLYFTDFVGEAVFHGDPHTAAPNSEAALNYEAGVRFGCWGMIAVAFSC 453
QY 336 LVFSLVMDRLVORFGTRAVYLASVAAPV---AAGATCLSHSVAVVVTASALTGTFTSAL 392
Db 454 SIYSLSVTKLMKMGFTGKAVYISGMIVYGLMVLGMLPDKGVLVFTSA---GILYGTI 510
QY 393 QILPYTLASLYHREKQVFLPKYRGDTGASSEDLSMTSFLPGPKGAPFPNGHVGAGSG 452
Db 511 FTVPPFLVARYH-AKNCFSIK-----NG-----532
QY 453 LLPPPALCGASACDVSVRVVVGTEPTEARVVP---CGICLDLAILDSAFLLSVAFLS 509
Db 533 -----BIVPLKQARGIGTDAIISVMVFFIAQLIVSLUS 564
QY 510 MGSIVQLSOSVTAYMVVSAAGLGLVAIFYATQVVF 543
Db 565 VGPLVSWMDTTCAVLVYASTFSLFLAIAAMFVLV 598

RESULT 7
QVR110 PRELIMINARY; PRT; 112 AA.
AC QVR110
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 11.4 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024519; AAH24519.1; -
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 112 AA; 11393 MW; B9D9DE79FEF09FB7 CRC64;

```



```
QY 477 PTEARVPGRGICLDLAILDSAPLLSQVAPSLFMGSIQVLSQSVTAYMVSAAAGLGLVAIY 536
Db 532 ---AQIIVGAG-----LGLVNRAGSVIVVLSASSISLIGCI 566
QY 537 F 537
Db 567 F 567

RESULT 10
Q8K4S3 PRELIMINARY; PRT; 751 AA.
AC Q8K4S3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Proton-associated sugar transporter A.
GN PAST-A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Shimokawa N., Okada J., Haglund K., Dikic I., Koibuchi N., Miura M.;
RT "Past-A, a new class proton-associated sugar transport facilitator in
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB075229; BAB97313.1; -.
KW Sugar transport.
SQ SEQUENCE 751 AA; 81751 MW; 5C33E1C2E4A1B2EC CRC64;

Query Match 15.6%; Score 446; DB 11; Length 751;
Best Local Similarity 22.3%; Pred. No. 1.3e-25;
Matches 160; Conservative 79; Mismatches 206; Indels 272; Gaps 15;

QY 17 QLLVNLTLFGLVCLAGITYPPPLLEVGVEKFWMTWILGIPVLGVCVPLIGASD 76
Db 86 ELLFNGCILFGIEFSYAMETAYTPVLLQMLPDQLSYLWFSPIGLFLQLLQAGWSD 145
QY 77 HWGRGRRRPPFTWALSGLTLLSLFLPRAGLWAGLLCPDP---RPLEALLTLGVGLLD 133
Db 146 RCTSRGRRRPPFLVLAIGALLGLSLLN-GRDIGHALADATNHWGILLITVCGVLM 204
QY 134 FCQGVCTPLEALLSDFRDPHCRQAYSVYAFMISLGGCLGYLLPAIDWDTSAIYLG 193
Db 205 FSADSADNPESHAYMMDVCGPVDQR-GLNIHALMAGLGGFGYVVGIIHWKTSFGRALG 263
QY 194 TQECFLGLTLFLTCVAATLVA-----EEAA 222
Db 264 GOLRVY-IFTATLSTVTVTLVSPERPLRPLGKRTAMKSPSLPPLPPVLLLEGA 322
QY 223 LGTEP---AEGLSAPLSPHCCP----- 243
Db 323 -GUTLSTATSLYASFSSPSPPLTPKXGFSIRSDSLTGINFASFSTNIDSVL 381
QY 244 -----CRALAFNLGALL----- 257
Db 382 IDCFTAGHDNYLALPSSVPRQAI SVSPRAPDGYCOERLERREGPLTLGLDGVLRG 441
QY 258 -----PR----- 259
Db 442 SLDTSKPRASGILKRPQTALPDVAGNGPETSRRNRVTFSSQVANILLGVKYESLTG 501
QY 260 -----LHOLCC---RMPTLRRLFAELCSNWMALMTFLFTYDFVGEGLYQGVPR 306
Db 502 SSSQSQPLSLRLCSTIYNMPPVRLNWHFLGLMSFEGMLLFTYDFMGVGVFQDQPK 561
QY 307 AEPGTEARRHYDEGVGRMSGLFLQCAISLVFSLVMDRLVQRGTRAVYLASVAAPVAA 366
Db 562 APHASEAYQKNSVGTWCGMCMCIYAFSAFYSAILEKLECLSVRTLFIYVLLFGLGT 621
```

```
QY 367 GATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGASSED 426
Db 622 GLATLSRNLYVVLSCTHYGILFSLTCLTLPYSLCDYYQSK-----KFAQ-----SSADG 671
QY 427 LMTSFLPGPKGAPPPNGHVAGGGLLPPPPALCGASACDVSVRVVVGEPTARVVPGR 486
Db 672 TR-----R 674
QY 487 GICLDLAILDSAPLLSQVAPSLFMGSIQVLSQSVTAYMVSAAAGLGLVAIYATQVVF 543
Db 675 GMGVDISLLSCQYFLAQILVSLVLP-----LTSVAGSANGVMYFASLVSP 720

RESULT 11
Q8MMT9 PRELIMINARY; PRT; 754 AA.
AC Q8MMT9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SUCI-sucrose proton symporter.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungstad B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AC117076; AAM33692.1; -.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 754 AA; 82774 MW; 614C473C832ACCE4 CRC64;

Query Match 13.8%; Score 394.5; DB 5; Length 754;
Best Local Similarity 23.5%; Pred. No. 1.1e-21;
Matches 128; Conservative 90; Mismatches 200; Indels 127; Gaps 13;

QY 13 HRKA---QLLVNLTLFGLVCLAGITYPPPLLEVGVEKFWMTWILGIPVLGVCV 68
Db 197 HKRTLPFLNLCITCLFGLVGFQWALQIAFSTPLFLELGEVQKWSYIWLAGISGLIVQ 256
QY 69 PLLGSASDHWGRYGRRRPFIWALSGLILLSLFLPRAGLWAGLL--CPDPRPLEALLI 126
Db 257 PLVGVTDRSECRFGRKPFILIGSVISGLVLSNAETFGSYFGDSQKKSIAISIAI 316
QY 127 LGVGLDPCQGVCTPLEALLSDFRDPHCRQAYSVYAFMISLGGCLGYLLPAIDWDT 186
Db 317 VGFWILDNNNAVQAPCRALLVDI-AAPSQSLGSLFSLMLGTGNLLGYMMSI--DLV 373
QY 187 ALAPYLGTQEECLFGLTLFLTCVAATL-LVAEEAALGPTPEAGLSAPSLSPHCCP 245
Db 374 RMVPMKKTTRALFTLSIMVLLFCVVMTLGFVTEQYIRVND-QSVENP----- 422
QY 246 ARLAFNLGALLPRLHQLCCMRPTRLRLFAELCSNWMALMTFLFTYDFVGEGLYQGV 305
Db 423 -----LKTMPKGIVKMPTYLQRLCAVQFFSWIGFSLVFTITWGVNVFGDP 471
QY 306 RAEPGTEARRHYDEGVGRMSGLFLQCAISLVFSLVMDRLVQRGTRAVYLASVAAPV 365
Db 472 NAEYSDSRILFQDGVGRMSGLSLITISSGITIAVSLIPLFLVKFIDMKYIYIGNL----- 526
QY 366 AGATCLSHSVAVVTAS-----AALTGTFSSALQILPYTLASLYHREKQVFLPKYRG 418
Db 527 --LQCIFFALFYVESKISGLILLIASTGIPWALVMTLPFSIVGM----- 568
QY 419 CGASSEDLSMTSFLPGPKGAPPPNGHVAGGGLLPPPPALCGASACDVSVRVVVGEPT 478
```

```
Db 569 -GVEDNES-----SG-----577
QY 479 EARVWPGRGICLDLAILDLSAFLLSOVAPSLFMSGIYQLSQSVTAYMVSAAGLGLVAIYPA 538
Db 578 -----LNIITLNIFFVVPQMVVSLIGLILDLSSKGNVYSLT---GSVASFFA 623
QY 539 TQVVF 543
Db 624 TLFCF 628

RESULT 12
O65803 PRELIMINARY; PRT; 515 AA.
AC O65803;
DT 01-AUG-1998 (T-EMBLrel. 07, Created)
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Sucrose/H+ symporter.
GN SUT2.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nantaise;
RX MEDLINE=99063785; PubMed=9847123;
RA Shaky R.; Sturm A.;
RT "Characterization of source- and sink-specific sucrose/H+ symporters
    from carrot.";
RL Plant Physiol. 118:1473-1480 (1998).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; Y16768; CAA76369.1; -.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; sugar tr; 1.
DR TIGRfam; TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
SQ SEQUENCE 515 AA; 54424 MW; 897B90657C9E243C CRC64;

Query Match 12.1%; Score 347.5; DB 10; Length 515;
Best Local Similarity 25.9%; Pred. No. 2.7e-18;
Matches 131; Conservative 80; Mismatches 215; Indels 79; Gaps 17;
QY 17 QLLVNLTLTFGLEVCCLAAGITYVPPILLLEVGVEKEFTMWLGIGPVGLVGLVCPPLGASD 76
Db 34 KLVLVAIAAGVQFGWALQSLTTPYVQLLGI PHKWAAYIWLGCPISGMLVQPIVGYSD 93
QY 77 HWRGRYRRRPFTWALSGLILLSLFLIPRAGWLAGL-----LCPDPRPLELALLIGV 129
Db 94 HCQSFGRRRPFTASGAGCVAISVILI---GFAADISYKAGDDMSKTLKPRAVTVFVIGF 150
QY 130 GLLDPCGQVCFTEPLEALLSLFR-DPDHCRQAYSVVAFMISLGGCLGY-----LL 178
Db 151 WILDVANNMLOGPCRRALLADLCSDGTRMRSAFYSFFMAVGNILGYAAGSYNNLYKLF 210
QY 179 PAIDWDTALAPYLGTECECLFGLLTLITLTCVAATLLVAEEAALGPTPEAGLSAPSLS 238
Db 211 PFS--KTHACDLYCANLKSCFTIISALLIITVALSVVRENS--GPPDDAABEEP--- 263
QY 239 PHCCPCRLAFRNIGALLPRHLQCCRPRTLRLFVAELCSWMALMTFTLYTDFVGE 298
Db 264 ----PSSGKIPV--FGEILGALKOL----PRMILLIITCLNIAWTAFFILFDTDMGR 313
QY 299 GLYQGVPRAPGTEARRHYDEGVRMGSLGLFQCAISLVFSLVMDRLVQRFTRAVYLAS 358
Db 314 EYGGT--AGQG-----KLVDQGVRAAGALLNLSVVLGTSIAVEYLVRGVGVKTI-LWG 366
QY 359 VAAPFVAAGATCL-----SHSV-----AVVTASA-----ALTGTFFSALQIL 395

Query Match 12.1%; Score 347.5; DB 10; Length 515;
Best Local Similarity 25.9%; Pred. No. 2.7e-18;
Matches 131; Conservative 80; Mismatches 215; Indels 79; Gaps 17;
QY 17 QLLVNLTLTFGLEVCCLAAGITYVPPILLLEVGVEKEFTMWLGIGPVGLVGLVCPPLGASD 76
Db 34 KLVLVAIAAGVQFGWALQSLTTPYVQLLGI PHKWAAYIWLGCPISGMLVQPIVGYSD 93
QY 77 HWRGRYRRRPFTWALSGLILLSLFLIPRAGWLAGL-----LCPDPRPLELALLIGV 129
Db 94 HCQSFGRRRPFTASGAGCVAISVILI---GFAADISYKAGDDMSKTLKPRAVTVFVIGF 150
QY 130 GLLDPCGQVCFTEPLEALLSLFR-DPDHCRQAYSVVAFMISLGGCLGY-----LL 178
Db 151 WILDVANNMLOGPCRRALLADLCSDGTRMRSAFYSFFMAVGNILGYAAGSYNNLYKLF 210
QY 179 PAIDWDTALAPYLGTECECLFGLLTLITLTCVAATLLVAEEAALGPTPEAGLSAPSLS 238
Db 211 PFS--KTHACDLYCANLKSCFTIISALLIITVALSVVRENS--GPPDDAABEEP--- 263
QY 239 PHCCPCRLAFRNIGALLPRHLQCCRPRTLRLFVAELCSWMALMTFTLYTDFVGE 298
Db 264 ----PSSGKIPV--FGEILGALKOL----PRMILLIITCLNIAWTAFFILFDTDMGR 313
QY 299 GLYQGVPRAPGTEARRHYDEGVRMGSLGLFQCAISLVFSLVMDRLVQRFTRAVYLAS 358
Db 314 EYGGT--AGQG-----KLVDQGVRAAGALLNLSVVLGTSIAVEYLVRGVGVKTI-LWG 366
QY 359 VAAPFVAAGATCL-----SHSV-----AVVTASA-----ALTGTFFSALQIL 395
```

```

QY 359 VAAPFAAGATCL-----SHSV-----AVVTASA-----ALTGFTFSALQIL 395
Db 367 FNFILAIAGLMTVVVSKVAQHOREHSANGQLLPSPAGVAGALSFLSILGPLSTYSI 426
QY 396 PYTLASLYHREKQVFLPKYRGDTGGASSEDLSMTSLPQPKPG-----APFPNGHVGA 448
Db 427 PFALASIYSGSGAGGGLSLGVNLAIIVPQMIVSVLAGPFDLSFGGGLNLPAPVGAISA 486
QY 449 GSGGLLP-----PPPALCGASACDVS 469
Db 487 AISGVLAIVLLPKPSKDAKSLSL 511

RESULT 14
Q9SQK6 PRELIMINARY; PRT; 501 AA.
AC Q9SQK6;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Putative sucrose transporter.
GN VVSUC11.
OS Vitis vinifera (Grape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC Vitis.
OX NCBI_TaxID=29760;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Shiraz; TISSUE=Fruit;
RA Davies C., Wolf T., Robinson S.P.;
RT "Three putative sucrose transporters are differentially expressed in
RT grapevine tissues."
RL Plant Sci. 147:93-100(1999).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL: AF021808; AAF08329.1; -.
DR InterPro: IPR005828; Sub transporter.
DR InterPro: IPR005989; Suc/H_symport.
DR Pfam: PF00083; sugar tr; 1.
DR TIGRFAMs: TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
SQ SEQUENCE 501 AA; 53938 MW; 4D0D4DE2EF2F4BA8 CRC64;

```

```

Query Match 12.1%; Score 346.5; DB 10; Length 501;
Best Local Similarity 27.0%; Pred. No. 3.1e-18;
Matches 153; Conservative 83; Mismatches 197; Indels 133; Gaps 22;

QY 17 QLLLVNLLTFGLVCLAAAGITVVPPLLELVGVVEKFTMTVLGIGPVGLVCVPLIGSAD 76
Db 29 RLURASVACGIQFGWALQSLTTPVQELGIPHAMSSIIWLCPLSGLLVQPLVGLSD 88
QY 77 HWRGRRRRPFIFWALSGLLSLFLIPRAGWLAGLL--CPDPRPLEALLILGVGLDF 134
Db 89 RCNSRFRGRRPFIVAGATSIWVAVLIIGFSADIGLLGDGADRPRPRAVATFVVGFWLLDV 148
QY 135 CGQVCTFPLEALLSDLPDPH--CQAYSVYAFMISLGCGLYLLPA-----IDWDS 186
Db 149 ANNVTOGPCRALLADL--TEKHRRTRVANYAFSLFIAVGNVLGATGYSYGNWFRIFWFTS 207
QY 187 ALAPYLGTQECU-----FGLLTILFTCVAAATLLVAEEAALGPTPEAGLSAPLSPHCC 242
Db 208 TSS-----CNADCANLKSFAFLDIIF--IAITYISITAA-----QELPLSSSSRTHIS 255
QY 243 -----PCRAFLAPRNILGALLPRHLQCCMRPTRLR-----LFVAELCSWMAWMTFT 289
Db 256 EEMAESTHAQEA-----LWELF-----GTLRYLSGSIWILFVTAL--TWIGLLPFL 301
QY 290 LFTYDFVGEGLYQGVPAEPCGTARHRYDEGVMSGLGLFQCAISLVFSLVMDRLVQRF 349
Db 302 LFTDWMGRIYGGKP-----NEGQNTYGVRMGALGMLNSVILGITSVLMEKLCRW 355
QY 350 GTRAVVLASVAAPFAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQV 409

```

```

Db 356 GAGFVW-----GLSNILMSLCFLMLLILS---AVVGHMD--- 386
QY 410 FLPKYRGDTGGASSEDLSMTSLPQPKCAPPNHGVGAGSGGLLPPPPALCGASACDVS 469
Db 387 -----FLGHDLP-----PSGVWIA-----ALIVFSILGIP 411
QY 470 VRVVVGEV---TEARVVP--GRGICLDLALDLSAFLLSOPVAPSLFMGSIQVL-----SQSV 520
Db 412 LAITYSVVVALISTRIESLGLQGLSMGVNLAIIVPQIVNSLGSGLPQDLFGGNSPSL 471
QY 521 TAYMVSAAAGLGLVAIYFATQVVFDS 546
Db 472 AVAAVAAPASGLVAIILAIIPRSSADKS 497

RESULT 15
Q9SLN7 PRELIMINARY; PRT; 515 AA.
AC Q9SLN7;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Sucrose transporter protein.
GN CSUT.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root;
RA Wu N., Diao F., Zhang L., Huang M.;
RT "Sucrose transporter protein."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL: AB036758; BAA89458.1; -.
DR InterPro: IPR005828; Sub transporter.
DR InterPro: IPR005989; Suc/H_symport.
DR Pfam: PF00083; sugar tr; 1.
DR TIGRFAMs: TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
SQ SEQUENCE 515 AA; 54523 MW; 2B548D1D9DFF51AC CRC64;

```

```

Query Match 12.1%; Score 346.5; DB 10; Length 515;
Best Local Similarity 27.2%; Pred. No. 3.2e-18;
Matches 137; Conservative 76; Mismatches 215; Indels 75; Gaps 19;

QY 17 QLLLVNLLTFGLVCLAAAGITVVPPLLELVGVVEKFTMTVLGIGPVGLVCVPLIGSAD 76
Db 34 KULVVAATAAGVQFGWALQSLTTPVQELGIPHWAAVYIWLCPISGMLVQPIVGYSD 93
QY 77 HWRGRRRRPFIFWALSGLLSLFLIPRA---GMLAG--LLCPDPRPLEALLILGVGLL 132
Db 94 HCQSSFRGRRPFIASGAGCAISVILIGPAADIGYKAGDMSKTLKPRAVTFVIGFWIL 153
QY 133 DFCQVCTFPLEALLSDLPF--PPDCHCRQAYSVYAFMISLGGCLGY-----LLPAI 181
Db 154 DVANNMLOGPCRALLADLCSGDTRRMRSSANAFYSFPMVGNILGYAAGSYNNLYKLFPFS 213
QY 182 DWDTSALAPYLGTQECU--LFGLLTILFTCVAAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 214 ---KTHACDLYCANLKSCFIISIALIITVVA--LSVREKOWSP--DDADADEP----- 263
QY 241 CCPCRAFLAPRNILGALLPRHLQCCMRPTRLRRLFVAELCSWMAWMTFTLFTYDFVGEGL 300
Db 264 --PSSGKIPV--FGELLGALKDL---PREMLLLIIVTCLNWIAMPFFILFDTDMGREGI 315
QY 301 YQGVPAEPCGTARHRYDEGVMSGLGLFQCAISLVFSLVMDRLVQRFCTRAVVLASVA 360
Db 316 YGGT--AGKG-----KLYDQGVRAAGSLGLLNSVILGITSVIAEYLVRGVGGVKI--LWGVV 368
QY 361 AFPVAAGATCL-----SHSV-----AVVTASA-----ALTGFTFSALQILPY 397

```

359	DB	NFILAIGLVMTVVSVKVAHQREHSANGQLLPSAGVKGALSLFSLGIPUSITYSIPF	428
398	QY	TLASLYHREKQVFLPKYRGDTGGASSEDLSMTSFLPCKPG-----APFPNGHVGAGG	450
429	DB	ALASTYSSGSGAGQGLSLGLVNLAIVVPQIMVSLAGPDSLFGGGLNLPFAVVGALSAIAI	488
451	QY	SGLLP-----PPPALCGASACDVS	469
489	DB	SGVLAIVLLPKPSKDAASKLSLS	511

Search completed: December 3, 2003, 17:20:41
Job time : 42 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 17:19:21 ; Search time 47 seconds
(without alignments)
1867.570 Million cell updates/sec

Title: US-09-593-793a-113

Perfect score: 553

Sequence: 1 MVQRLVMSRLRHKQALL.....AIYFATQVDFKSLAKISA 553

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	553	100.0	553	19	Amino acid encoded
2	553	100.0	553	19	Prostate tumour sp
3	553	100.0	553	21	Protein encoded by
4	553	100.0	553	21	Human immunogenic
5	553	100.0	553	22	Human prostate cDN
6	553	100.0	553	22	Human prostate tum
7	553	100.0	553	22	Human prostate-spe
8	553	100.0	553	22	Human prostate-spe
9	553	100.0	553	22	Human P501S invent

10	553	100.0	553	22	AA774800	Prostate tumour an
11	553	100.0	553	23	ABG94411	Human prostate tum
12	553	100.0	553	23	ABG76665	Prostate tumour pr
13	553	100.0	553	23	ABG77575	Human mast cell re
14	553	100.0	553	23	ABG61900	Prostate cancer-as
15	553	100.0	553	23	ABG52222	Human L1-12 protei
16	553	100.0	553	23	AAU10324	Human PROST 03. H
17	553	100.0	553	23	AAU82643	Human breast tumou
18	553	100.0	553	23	ABU71653	Prostate cancer ap
19	553	100.0	1079	22	ABU74830	Prostate tumour an
20	553	100.0	1079	24	ABU71860	Prostate specific
21	501	90.6	710	23	AAW50661	Thioredoxin-ubiqui
22	499	90.2	595	22	AAW01318	Alpha prepro-P501S
23	320	57.9	359	24	ABU71887	Human prostate spe
24	314	56.8	530	23	AAW50662	Human prostate spe
25	263	47.6	371	22	AAU69875	Thioredoxin-ubiqui
26	263	47.6	371	22	AAW01230	Human prostate cDN
27	263	47.6	371	22	AAW01362	P53S splice varia
28	263	47.6	371	23	ABG64105	Human albumin fusi
29	263	47.6	371	23	ABG95335	Human P553S splice
30	263	47.6	371	24	ABU71766	Prostate cancer as
31	263	47.6	400	22	AAU69907	Human prostate pro
32	263	47.6	400	22	AAW01262	Human prostate-E2 cons
33	263	47.6	400	23	ABG95367	Rai2-P501S-E2 cons
34	263	47.6	400	24	ABU71798	Prostate cancer as
35	255	46.1	255	20	AAW85068	Protein encoded by
36	255	46.1	255	21	AAW29268	Human prostate-rel
37	255	46.1	255	22	AAU04205	Prostate-specific
38	255	46.1	255	23	AAO19084	Human prostate-spe
39	206	37.3	231	21	AAW54369	Amino acid sequenc
40	187	33.8	326	23	ABW77571	Human mast cell re
41	161	29.1	252	22	AAE01423	Human secreted pro
42	87	15.7	123	22	AAU69873	Human prostate cDN
43	87	15.7	123	22	AAW01228	P553S splice varia
44	87	15.7	123	23	ABG95333	Human P553S splice
45	87	15.7	123	24	ABU71764	Prostate cancer as

ALIGNMENTS

RESULT 1

AAW71869

ID AAW71869 standard; Protein; 553 AA.

AC AAW71869;

06-JAN-1999 (first entry)

Amino acid encoded by prostate tumour clone L1-12.

Prostate; cancer; tumour; vaccine; immunogen; clone.

Homo sapiens.

WO9837093-A2.

27-AUG-1998.

25-FEB-1998; 98WO-US03492.

09-FEB-1998; 98US-0020956.

25-FEB-1997; 97US-0806099.

01-AUG-1997; 97US-0904804.

(CORI-) CORIXA CORP.

Dillon DC, Xu J;

WPI, 1998-609886/51.

DR N-PSDB; AAV61201.

Polypeptides comprising immunogenic portions of prostate proteins -

PT used in a vaccine for the treatment of prostate cancer
 XX Example 1; Page 82-84; 130pp; English.
 XX

CC The present sequence is an immunogenic portion of a prostate tumour
 CC protein. The immunogen, or the DNA encoding it, can be used as a
 CC vaccine for the treatment of prostate cancer. The immunogen was
 CC isolated from a prostate tumour cDNA library obtained by subtracting
 CC a prostate tumour cDNA expression library with a normal tissue cDNA
 CC library.
 XX

SQ Sequence 553 AA;

Query Match 100.0%; Score 553; DB 19; Length 553;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVORLWVSRLLRHRKAQQLLVNLLTFGLEVCLAAGITYVPPLLELVGVBEKFTMWLGIG 60
 DB 1 MVORLWVSRLLRHRKAQQLLVNLLTFGLEVCLAAGITYVPPLLELVGVBEKFTMWLGIG 60
 QY 61 PVGLVCVPLLGASDHWGRGRRRPFIFWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
 DB 61 PVGLVCVPLLGASDHWGRGRRRPFIFWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
 QY 121 ELALLILGVGLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVGLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDWDTSLAPYLGTQBECLFGLLTFLITCVAAATLLVAEEAALGPTPEAGLSAPLSPH 240
 DB 181 IDWDTSLAPYLGTQBECLFGLLTFLITCVAAATLLVAEEAALGPTPEAGLSAPLSPH 240
 QY 241 CCPCARLAFRNLGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLYTDFVGEGL 300
 DB 241 CCPCARLAFRNLGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLYTDFVGEGL 300
 QY 301 YQGVPAEPTGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 DB 301 YQGVPAEPTGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 QY 361 AFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 AFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVVGEPTA 480
 DB 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVVGEPTA 480
 QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVMSAAGLGLVAYIFATQ 540
 DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVMSAAGLGLVAYIFATQ 540

RESULT 2

AAW69385
 ID AAW69385 standard; Protein; 553 AA.

XX AAW69385;

XX 25-MAR-2003 (updated)
 DT 08-DEC-1998 (first entry)

XX Prostate tumour specific gene clone L1-12 protein.

XX Prostate tumour specific gene; human; prostate cancer; detection;
 KW therapy.
 XX

OS Homo sapiens.

XX WO9837418-A2.
 PN 27-AUG-1998.
 XX 25-FEB-1998; 98WO-US03690.
 XX 25-FEB-1997; 97US-0806596.
 PR 01-AUG-1997; 97US-0904809.
 PR 09-FEB-1998; 98US-0020747.
 XX (CORI-) CORIXA CORP.
 XX Dillon DC, Xu J;
 PI WPI; 1998-480805/41.
 XX N-PSDB; AAV58586.
 DR Novel human prostate specific tumour protein and fragments - useful
 PT for detecting and treating prostate cancer
 XX Example 1; Page 87-89; 141pp; English.
 PS This sequence is encoded by a human prostate tumour specific gene, and
 CC can be used in the method of the invention. The method is for detecting
 CC prostate cancer comprising contacting a biological sample with an agent
 CC able to bind an immunogenic portion of a prostate protein (such as
 CC this protein sequence). An antibody which binds to an immunogenic
 CC portion of the prostate protein, and the method can be used to detect;
 CC monitor progression of, or treat prostate cancers. The antibody may
 CC also be conjugated to a therapeutic agent for use in therapy of prostate
 CC cancers.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 XX

SQ Sequence 553 AA;

Query Match 100.0%; Score 553; DB 19; Length 553;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVORLWVSRLLRHRKAQQLLVNLLTFGLEVCLAAGITYVPPLLELVGVBEKFTMWLGIG 60
 DB 1 MVORLWVSRLLRHRKAQQLLVNLLTFGLEVCLAAGITYVPPLLELVGVBEKFTMWLGIG 60
 QY 61 PVGLVCVPLLGASDHWGRGRRRPFIFWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
 DB 61 PVGLVCVPLLGASDHWGRGRRRPFIFWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
 QY 121 ELALLILGVGLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVGLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDWDTSLAPYLGTQBECLFGLLTFLITCVAAATLLVAEEAALGPTPEAGLSAPLSPH 240
 DB 181 IDWDTSLAPYLGTQBECLFGLLTFLITCVAAATLLVAEEAALGPTPEAGLSAPLSPH 240
 QY 241 CCPCARLAFRNLGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLYTDFVGEGL 300
 DB 241 CCPCARLAFRNLGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLYTDFVGEGL 300
 QY 301 YQGVPAEPTGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 DB 301 YQGVPAEPTGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 QY 361 AFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 AFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVVGEPTA 480
 DB 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVVGEPTA 480
 QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVMSAAGLGLVAYIFATQ 540

CC the polypeptides, antibodies against the polypeptides and vaccines
 CC comprising them can be used for inhibiting the development of prostate
 CC cancer in a patient. The polypeptides can be used to generate antibodies
 CC or anti-idiotypic antibodies for passive immuno therapy. A portion of
 CC the polynucleotides encoding the polypeptides can be used as a probe or
 CC to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
 CC AAY82000 to AAY82020 represent sequences used in the exemplification of
 CC the present invention.

XX
 SQ Sequence 553 AA;

Query Match 100.0%; Score 553; DB 21; Length 553;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLLVGVGVEKFTMTVLGIG 60
 DB 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLLVGVGVEKFTMTVLGIG 60

QY 61 PVLGLVCPVLLGSASDHWGRGRRRPFITWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
 DB 61 PVLGLVCPVLLGSASDHWGRGRRRPFITWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLILVGLLDFCGQVCFTEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILVGLLDFCGQVCFTEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTSALAPYLGTQECFLGILLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLPH 240
 DB 181 IDWDTSALAPYLGTQECFLGILLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLPH 240

QY 241 CCPCRARLAFRNLGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLYTDFVGEGL 300
 DB 241 CCPCRARLAFRNLGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLYTDFVGEGL 300

QY 301 YQGVPAEPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFTRAVYLASVA 360
 DB 301 YQGVPAEPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFTRAVYLASVA 360

QY 361 AFPVAAGATCLSHSVAVVTASAALTGFTFSAQLPYTLASLVHREKQVFLPKYRGDTGG 420
 DB 361 AFPVAAGATCLSHSVAVVTASAALTGFTFSAQLPYTLASLVHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGKPGAPFNGHVAGGSGLLPPPPALCGASACDVSRVVVGEPTA 480
 DB 421 ASSEDSLMTSFLPGKPGAPFNGHVAGGSGLLPPPPALCGASACDVSRVVVGEPTA 480

QY 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMSIVQLSQSVTAYMWSAAGLGLVAYIPATQ 540
 DB 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMSIVQLSQSVTAYMWSAAGLGLVAYIPATQ 540

QY 541 VFDFKSDLAKYSA 553
 DB 541 VFDFKSDLAKYSA 553

RESULT 5
 AAU69763
 ID AAU69763 standard; Protein; 553 AA.
 XX
 AC AAU69763;
 XX
 AC
 XX
 DT 30-JAN-2002 (first entry)
 XX
 DE Human prostate cDNA encoded protein #3.
 XX
 DE Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN W0200173032-A2.
 XX
 XX
 PD 04-OCT-2001.

XX
 PF 27-MAR-2001; 2001WO-US09919.
 XX
 PR 27-MAR-2000; 2000US-0536857.
 PR 09-MAY-2000; 2000US-0568100.
 PR 12-MAY-2000; 2000US-0570737.
 PR 13-JUN-2000; 2000US-0593793.
 PR 17-JUN-2000; 2000US-0605783.
 PR 10-AUG-2000; 2000US-0636215.
 PR 29-AUG-2000; 2000US-0651236.
 PR 06-SEP-2000; 2000US-0657279.
 PR 02-OCT-2000; 2000US-0679426.
 PR 10-OCT-2000; 2000US-0685166.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedwick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
 XX
 DR WPI; 2001-639232/73.
 DR N-PSDB; AAS63557.
 XX
 PT New human prostate-specific polypeptides and polynucleotides useful for
 PT the diagnosis and treatment of cancer, especially prostate cancer -
 XX
 PS Claim 2; Page 269-270; 579pp; English.
 XX
 CC The invention relates to isolated prostate-specific
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
 CC antibodies raised against the polypeptides for antigenic epitopes
 CC derived from them) and antigen-presenting cells expressing the
 CC polypeptides. The antibodies are useful for detecting the presence of
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
 CC the antigen-presenting cells are useful for stimulating and/or expanding
 CC T cells specific for a tumour protein, and for inhibiting the development
 CC of cancer especially prostate cancer. Compositions comprising an immune
 CC polynucleotide and/or polypeptide are useful for stimulating an immune
 CC response, and for treating cancer. The oligonucleotide is useful for
 CC detecting cancer. The present sequence is a prostate specific
 CC polypeptide of the invention.
 XX
 SQ Sequence 553 AA;

Query Match 100.0%; Score 553; DB 22; Length 553;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLLVGVGVEKFTMTVLGIG 60
 DB 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLLVGVGVEKFTMTVLGIG 60

QY 61 PVLGLVCPVLLGSASDHWGRGRRRPFITWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
 DB 61 PVLGLVCPVLLGSASDHWGRGRRRPFITWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLILVGLLDFCGQVCFTEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILVGLLDFCGQVCFTEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTSALAPYLGTQECFLGILLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLPH 240
 DB 181 IDWDTSALAPYLGTQECFLGILLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLPH 240

QY 241 CCPCRARLAFRNLGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLYTDFVGEGL 300
 DB 241 CCPCRARLAFRNLGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLYTDFVGEGL 300

QY 301 YQGVPAEPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFTRAVYLASVA 360
 DB 301 YQGVPAEPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFTRAVYLASVA 360

QY 361 AFPVAAGATCLSHSVAVVTASAALTGFTFSAQLPYTLASLVHREKQVFLPKYRGDTGG 420
 DB 361 AFPVAAGATCLSHSVAVVTASAALTGFTFSAQLPYTLASLVHREKQVFLPKYRGDTGG 420

CC (I) and the antibodies are also used in the detection of cancer in a
CC patient. The cancer that is diagnosed or treated is particularly
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC (I) can be used for monitoring the progression of cancer in a patient.
CC (I) and (II) can also be used to improve diagnostic and therapeutic
CC methods for prostate cancer. They can indicate the level of metastasis
CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to
CC AAH01318 represent polynucleotide and amino acid sequences used in the
CC exemplification of the present invention.

XX Sequence 553 AA;
SQ Query Match 100.0%; Score 553; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLVWSRLRRHKAQLLVNLLTFGLVCLAAAGITYVPPLLLEVGVEKFTMWLGIG 60
Db 1 MVQRLVWSRLRRHKAQLLVNLLTFGLVCLAAAGITYVPPLLLEVGVEKFTMWLGIG 60

Qy 61 PVGLVVCVPLLGASDHWGRGRRRPFIIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
Db 61 PVGLVVCVPLLGASDHWGRGRRRPFIIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120

Qy 121 ELALLILGVCLDFCGOVCFPLEALLSDFRDPDHCROQSVYAFMISIGGCLGYLLPA 180
Db 121 ELALLILGVCLDFCGOVCFPLEALLSDFRDPDHCROQSVYAFMISIGGCLGYLLPA 180

Qy 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

Qy 241 CCPCRARLAFRNLAGLPRHLQCCRMPTLRLLFVAELCSWWMALMTFTFYDFVGEGL 300
Db 241 CCPCRARLAFRNLAGLPRHLQCCRMPTLRLLFVAELCSWWMALMTFTFYDFVGEGL 300

Qy 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRFTRAVYLASVA 360
Db 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRFTRAVYLASVA 360

Qy 361 APFVAAGATCLSHSVAVVTASAAALGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAAALGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420

Qy 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTEA 480

Qy 481 RVVPGRGICLDLALDSAPLLSOVAPSLFMGSIIVQLSQSVTAIVWSAAGLGLVAIYFATQ 540
Db 481 RVVPGRGICLDLALDSAPLLSOVAPSLFMGSIIVQLSQSVTAIVWSAAGLGLVAIYFATQ 540

Qy 541 VFEDKSLAKYSA 553
Db 541 VFEDKSLAKYSA 553

RESULT 8
AAG99002
ID AAG99002 standard; Protein; 553 AA.
XX
AC AAG99002;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human prostate-specific amino acid sequence L1-12/P501S.
XX
KW Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW chromosome 22q11.2; prostate-specific protein; chromosome 1;
KW prostate specific antigen; PSA.
XX
OS Homo sapiens.
XX

PN WO200134802-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30904.
XX
PR 12-NOV-1999; 99US-0439313.
PR 18-NOV-1999; 99US-0443686.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;
XX
DR WPI; 2001-308785/32.
XX
PT Isolated polypeptide comprising at least an immunogenic portion of a
PT prostate-specific protein, useful in the diagnosis and therapy of
PT prostate cancer -
XX
PS Claim 3; Page 167-168; 325pp; English.
XX
CC The present invention describes an isolated polypeptide (P1) comprising
CC at least an immunogenic portion of a prostate-specific protein, or its
CC variant. Also described are polynucleotides (N1) encoding (P1). (P1) and
CC (N1) have cytostatic activity and can be used in vaccine production.
CC The polypeptides, nucleic acids and antibodies from the present
CC invention are useful in the diagnosis and therapy of prostate cancer.
CC Prostate specific genes P704P, P712P, P775P, P775P and B305D are located
CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
CC region. Prostate specific antigen (PSA) P501S was located on
CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
CC polynucleotide and polypeptide sequences used in the exemplification
CC of the present invention.
XX
SQ Sequence 553 AA;
Query Match 100.0%; Score 553; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLVWSRLRRHKAQLLVNLLTFGLVCLAAAGITYVPPLLLEVGVEKFTMWLGIG 60
Db 1 MVQRLVWSRLRRHKAQLLVNLLTFGLVCLAAAGITYVPPLLLEVGVEKFTMWLGIG 60

Qy 61 PVGLVVCVPLLGASDHWGRGRRRPFIIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
Db 61 PVGLVVCVPLLGASDHWGRGRRRPFIIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120

Qy 121 ELALLILGVCLDFCGOVCFPLEALLSDFRDPDHCROQSVYAFMISIGGCLGYLLPA 180
Db 121 ELALLILGVCLDFCGOVCFPLEALLSDFRDPDHCROQSVYAFMISIGGCLGYLLPA 180

Qy 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

Qy 241 CCPCRARLAFRNLAGLPRHLQCCRMPTLRLLFVAELCSWWMALMTFTFYDFVGEGL 300
Db 241 CCPCRARLAFRNLAGLPRHLQCCRMPTLRLLFVAELCSWWMALMTFTFYDFVGEGL 300

Qy 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRFTRAVYLASVA 360
Db 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRFTRAVYLASVA 360

Qy 361 APFVAAGATCLSHSVAVVTASAAALGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAAALGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420

Qy 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTEA 480


```
CC AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences
CC used in the exemplification of the present invention.
XX
SQ Sequence 553 AA;
Query Match 100.0%; Score 553; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVBEKFTMWLGIG 60
Db 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVBEKFTMWLGIG 60
QY 61 PVGLVVCVPLLGASDHWRGRRPFTWALSIGLILSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVVCVPLLGASDHWRGRRPFTWALSIGLILSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWDTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
QY 241 CCPCRLARLAFRNIGALLPRLHQLCCMRPTRLRLFVAELCSWMLMTFTLYTDFVGEGL 300
Db 241 CCPCRLARLAFRNIGALLPRLHQLCCMRPTRLRLFVAELCSWMLMTFTLYTDFVGEGL 300
QY 301 YQGVPAEPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPAEPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
QY 361 AFPVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 AFPVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPFNHGVGAGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFNHGVGAGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553
RESULT 11
ABG94411
ID ABG94411 standard; Protein; 553 AA.
XX
AC ABG94411;
XX
DT 27-NOV-2002 (first entry)
XX.
XX Human prostate tumour protein partial sequence #3.
XX Human; immunogenic; prostate protein; prostate tumour protein;
XX prostate cancer; cytostatic; vaccine.
XX Homo sapiens.
XX
XX US2002090372-A1.
XX
XX 11-JUL-2002.
XX
XX 14-JUL-1998; 98US-0115453.
XX
XX 25-FEB-1997; 97US-0806099.
XX
XX 01-AUG-1997; 97US-0904804.
XX
```

```
PR 09-FEB-1998; 98US-0020956.
PR 25-FEB-1998; 98US-0030607.
XX
PA (XUJJ/) XU J.
PA (DILL/) DILLON D C.
XX
PI Xu J, Dillon DC;
XX
XX WPI: 2002-642373/69.
DR N-PSDB; ABS21254.
XX
PT Novel polypeptides useful as vaccines for inhibiting prostate cancer
development, comprise an immunogenic portion of prostate protein -
XX
XX Example 1; Page 59-61; 101pp; English.
XX
CC The present invention relates to a new polypeptide comprising an
immunogenic portion of a prostate protein. The invention is useful for
inhibiting the development of prostate cancer in a patient. The
invention is also useful as markers for diagnosing prostate cancer and
for monitoring diseases progression in patients. The present amino
acid sequence represents a human prostate tumour protein.
XX
SQ Sequence 553 AA;
Query Match 100.0%; Score 553; DB 23; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVBEKFTMWLGIG 60
Db 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVBEKFTMWLGIG 60
QY 61 PVGLVVCVPLLGASDHWRGRRPFTWALSIGLILSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVVCVPLLGASDHWRGRRPFTWALSIGLILSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWDTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
QY 241 CCPCRLARLAFRNIGALLPRLHQLCCMRPTRLRLFVAELCSWMLMTFTLYTDFVGEGL 300
Db 241 CCPCRLARLAFRNIGALLPRLHQLCCMRPTRLRLFVAELCSWMLMTFTLYTDFVGEGL 300
QY 301 YQGVPAEPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPAEPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
QY 361 AFPVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 AFPVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPFNHGVGAGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFNHGVGAGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIFYATQ 540
QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553
RESULT 12
ABG76665
ID ABG76665 standard; Protein; 553 AA.
```


CC for presenting information to identify the relative expression level of
 CC (I). (II) is used as a marker to detect, diagnose or identify an allergic
 CC response in a patient. The protein can also serve as a target that
 CC modulate gene expression or activity and as an antigen to raise
 CC polyclonal or monoclonal antibodies. (II) is useful for identifying
 CC agents that modulate expression of the protein or agents, such as
 CC agonists or antagonists. The agonists or antagonists are useful for
 CC modulating biological activity and function of (II) and thus are useful
 CC for alleviating disease conditions such as allergic hypersensitivity,
 CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis.
 XX
 SQ Sequence 553 AA;

Query Match 100.0%; Score 553; DB 23; Length 553;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLELVGVEKEFMTWVLGIG 60
 DB 1 MVQLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLELVGVEKEFMTWVLGIG 60

QY 61 PVGLVCVPLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
 DB 61 PVGLVCVPLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120

QY 121 ELALLILGVCLDFCGOVCFPLEALLSDLPDPDHCROAYSVYAFWISLGGCLGYLLPA 180
 DB 121 ELALLILGVCLDFCGOVCFPLEALLSDLPDPDHCROAYSVYAFWISLGGCLGYLLPA 180

QY 181 IDWDTSAALPYLTQBECLFGLLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
 DB 181 IDWDTSAALPYLTQBECLFGLLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240

QY 241 CCPCRLARLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLYTDFVGEGL 300
 DB 241 CCPCRLARLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLYTDFVGEGL 300

QY 301 YQGVPAEPCTEARRHYDEGVNGSLGLFLOCAISLVFSIWMRLVQRPGTRAVYLASVA 360
 DB 301 YQGVPAEPCTEARRHYDEGVNGSLGLFLOCAISLVFSIWMRLVQRPGTRAVYLASVA 360

QY 361 APVAAGATCLSHSVAVTASALTGTFTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 APVAAGATCLSHSVAVTASALTGTFTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
 DB 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480

QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFNGSIVQISQSVTATYVWSAAGLGLVAYIFATQ 540
 DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFNGSIVQISQSVTATYVWSAAGLGLVAYIFATQ 540

QY 541 VFDKSDLAKYSA 553
 DB 541 VFDKSDLAKYSA 553

RESULT 14
 ABG61900

ID ABG61900 standard; Protein; 553 AA.

XX ABG61900;

XX 15-AUG-2002 (first entry)

DT Prostate cancer-associated protein #101.

DE Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

XX Mammalia.

XX WO200230268-A2.

XX 18-APR-2002.

PD 12-OCT-2001; 2001WO-US32045.

XX 13-OCT-2000; 2000US-0687576.

PR 08-DEC-2000; 2000US-0733288.

PR 08-DEC-2000; 2000US-0733742.

PR 24-JAN-2001; 2001US-263957P.

PR 16-MAR-2001; 2001US-276791P.

PR 06-APR-2001; 2001US-281922P.

PR 24-APR-2001; 2001US-286214P.

PR 30-APR-2001; 2001US-0847046.

PR 04-MAY-2001; 2001US-288589P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
 WPI: 2002-471335/50.
 N-PSDB; ABK92217.

DR Detecting a prostate cancer-associated transcript in a cell in a
 XX patient, useful for diagnosing prostate cancer (PC) or screening
 PT modulators of PC, by determining if prostate cancer-associated genes
 PT are expressed in a prostate tissue -
 XX Claim 27; Page 386; 436pp; English.

PS The present invention relates to methods of detecting a prostate
 XX cancer-associated transcript in a cell from a patient. The method
 CC comprises contacting a biological sample from the patient with
 CC prostate cancer-associated polynucleotides (designated PC genes) that
 CC selectively hybridise to a sequence that is at least 80% identical
 CC to them. The prostate cancer-associated polynucleotide sequences
 CC are differentially expressed in prostate tumour tissue or in
 CC prostate cancer and are derived from the tissues of various
 CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
 CC The methods of the invention are useful for diagnosing and treating
 CC prostate cancer in mammals. The prostate cancer-associated genes are
 CC useful for diagnosing or treating prostate cancer, as well as for
 CC identifying modulators of prostate cancer or agents that inhibit
 CC prostate cancer. The nucleic acid sequences are particularly useful
 CC in gene therapy, as a vaccine or in antisense applications.
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins.
 XX Sequence 553 AA;

SQ Query Match 100.0%; Score 553; DB 23; Length 553;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLELVGVEKEFMTWVLGIG 60
 DB 1 MVQLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLELVGVEKEFMTWVLGIG 60

QY 61 PVGLVCVPLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
 DB 61 PVGLVCVPLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120

QY 121 ELALLILGVCLDFCGOVCFPLEALLSDLPDPDHCROAYSVYAFWISLGGCLGYLLPA 180
 DB 121 ELALLILGVCLDFCGOVCFPLEALLSDLPDPDHCROAYSVYAFWISLGGCLGYLLPA 180

QY 181 IDWDTSAALPYLTQBECLFGLLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
 DB 181 IDWDTSAALPYLTQBECLFGLLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240

QY 241 CCPCRLARLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLYTDFVGEGL 300
 DB 241 CCPCRLARLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTLYTDFVGEGL 300

QY 301 YQGVPRAPTEARRHYDEGVRMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db (DAYC/) STOLK J A.
PA (VEDV/) DAY C H.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
QY 361 APFAAGATCLSHSVAVVTASAAITGFTPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db (LISX/) LI S X.
PA (WANG/) WANG A.
PA (SKEI/) SKEIKY Y A W.
PA (HEPL/) HEPLER W T.
QY 421 ASSEDSLMTSFLPGPKGAPPPNGHVGAGGSGLLPPPPALCGASACDVSVRVVVGPTPEA 480
Db (HEND/) HENDERSON R A.
QY 421 ASSEDSLMTSFLPGPKGAPPPNGHVGAGGSGLLPPPPALCGASACDVSVRVVVGPTPEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVATQ 540
Db (KALO/) KALOS M D.
QY 541 VVFDKSDLAKYSA 553
Db (FANG/) FANGER G R.
541 VVFDKSDLAKYSA 553
(RETT/) RETTER M W.

RESULT 15

ABB95222 standard; Protein; 553 AA.

AC ABB95222;

DT 19-JUL-2002 (first entry)

DE Human L1-12 protein SEQ ID NO 113.

Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
gene therapy.

OS Homo sapiens.

PN US2002022248-A1.

PD 21-FEB-2002.

PF 12-JAN-2001; *2001US-0759143.

PR 25-FEB-1997; 97US-0806099.

PR 01-AUG-1997; 97US-0904804.

PR 09-FEB-1998; 98US-020956.

PR 25-FEB-1998; 98US-0030607.

PR 14-JUL-1998; 98US-0115453.

PR 23-SEP-1998; 98US-0159812.

PR 15-JAN-1999; 99US-0232149.

PR 09-APR-1999; 99US-0288946.

PR 13-JUL-1999; 99US-0352616.

PR 12-NOV-1999; 99US-0443686.

PR 18-NOV-1999; 99US-0443686.

PR 14-JAN-2000; 2000US-0483672.

PR 27-MAR-2000; 2000US-0536857.

PR 09-MAY-2000; 2000US-0568100.

PR 12-MAY-2000; 2000US-0570737.

PR 13-JUN-2000; 2000US-0593793.

PR 27-JUN-2000; 2000US-0605783.

PR 10-AUG-2000; 2000US-0636215.

PR 29-AUG-2000; 2000US-0651236.

PR 06-SEP-2000; 2000US-0657279.

PR 02-OCT-2000; 2000US-0679426.

PR 10-OCT-2000; 2000US-0685166.

XX (XUJJ/) XU J.

PA (DILL/) DILLON D C.

PA (MITC/) MITCHAM J L.

PA (HARL/) HARLOCKER S L.

PA (JIAN/) JIANG Y.

PA (KALO/) KALOS M D.

PA (FANG/) FANGER G R.

PA (RETT/) RETTER M W.

PA (STOL/) STOLK J A.
PA (DAYC/) DAY C H.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
PA (LISX/) LI S X.
PA (WANG/) WANG A.
PA (SKEI/) SKEIKY Y A W.
PA (HEPL/) HEPLER W T.
PA (HEND/) HENDERSON R A.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
DR WPI; 2002-255649/30.
XX
PT New prostate-specific polynucleotides for diagnosing and treating
PT diseases, in particular prostate cancer, and as markers for the
PT progression of cancer
XX
PS Claim 2; SEQ ID NO 113; 87pp; English.
XX
CC The present invention provides prostate-specific coding sequences and
CC their encoded proteins. These can be used in the diagnosis and treatment
CC of cancers, particularly prostate cancer. The present sequence is a
CC protein described in the invention.
XX
SQ Sequence 553 AA;

Query Match 100.0%; Score 553; DB 23; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHKAQQLLVNLLTFGLVCLAAAGTIVPPPLLLLVGVVEKEKMTWVLGIG 60
Db 1 MVQRLWVSRLLRHKAQQLLVNLLTFGLVCLAAAGTIVPPPLLLLVGVVEKEKMTWVLGIG 60
QY 61 PVLGLVCPVLLGASDHWGRYGRRRPFTWALSIGLLSLFLI PRAGWLAGLLCPDPRPL 120
Db 61 PVLGLVCPVLLGASDHWGRYGRRRPFTWALSIGLLSLFLI PRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLDFCGQVCFTPLEALLSDLPDPDHCQAYSVYAFMISLGLCYLLPA 180
Db 121 ELALLILGVGLDFCGQVCFTPLEALLSDLPDPDHCQAYSVYAFMISLGLCYLLPA 180
QY 181 IDWDTSAAPYLGTOECLFGLLLTLITCTVAATLLVAEEAALGPTBPAGLSAPLSLSPH 240
Db 181 IDWDTSAAPYLGTOECLFGLLLTLITCTVAATLLVAEEAALGPTBPAGLSAPLSLSPH 240
QY 241 CCPCARLAFRNLGALLPRLHQLCCMRPTRLRLFVAELCSMWALMTFTLFYTDVFGEL 300
Db 241 CCPCARLAFRNLGALLPRLHQLCCMRPTRLRLFVAELCSMWALMTFTLFYTDVFGEL 300
QY 301 YQGVPRAPTEARRHYDEGVRMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPRAPTEARRHYDEGVRMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
QY 361 APFAAGATCLSHSVAVVTASAAITGFTPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFAAGATCLSHSVAVVTASAAITGFTPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPPPNGHVGAGGSGLLPPPPALCGASACDVSVRVVVGPTPEA 480
Db 421 ASSEDSLMTSFLPGPKGAPPPNGHVGAGGSGLLPPPPALCGASACDVSVRVVVGPTPEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVATQ 540
QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

Search completed: December 3, 2003, 17:23:03
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 17:22:07 ; Search time 36 seconds
(without alignments)
2856.916 Million cell updates/sec

Title: US-09-593-793A-113

Perfect score: 553

Sequence: 1 MVQRLWVSRLLRHRKAQLL.....AIYFATQVVFVDSLAKYSA 553

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 684280 seqs, 185983659 residues

Word size : 0

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*

10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*

11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*

12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*

13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*

17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	553	100.0	553	9	US-09-745-288-101
2	553	100.0	553	9	US-09-838-785-2
3	553	100.0	553	9	US-09-759-143-113
4	553	100.0	553	9	US-09-780-669-113
5	553	100.0	553	9	US-09-030-606-113
6	553	100.0	553	9	US-09-822-827-113
7	553	100.0	553	9	US-09-115-453-113
8	553	100.0	553	10	US-09-232-880-113
9	553	100.0	553	10	US-09-895-793-113
10	553	100.0	553	10	US-09-895-814-113
11	553	100.0	553	12	US-10-144-678A-113
12	553	100.0	553	12	US-10-005-907-113
13	553	100.0	553	12	US-10-294-025-113
14	553	100.0	553	14	US-10-012-896-113
15	553	100.0	553	15	US-10-010-940-113

16	553	100.0	1079	9	US-09-822-827-947	Sequence 947, App
17	553	100.0	1079	10	US-09-895-793-947	Sequence 947, App
18	501	90.6	710	15	US-10-296-770-4	Sequence 4, Appli
19	320	57.9	359	9	US-09-822-827-974	Sequence 974, App
20	320	57.9	359	10	US-09-895-793-974	Sequence 974, App
21	314	56.8	530	15	US-10-296-770-5	Sequence 5, Appli
22	297	53.7	305	12	US-10-144-678A-1029	Sequence 1029, Ap
23	297	53.7	305	12	US-10-294-025-1029	Sequence 1029, Ap
24	263	47.6	371	9	US-09-759-143-708	Sequence 708, App
25	263	47.6	371	9	US-09-780-669-708	Sequence 708, App
26	263	47.6	371	9	US-09-822-827-708	Sequence 708, App
27	263	47.6	371	10	US-09-895-793-708	Sequence 708, App
28	263	47.6	371	10	US-09-895-814-708	Sequence 708, App
29	263	47.6	371	12	US-10-144-678A-708	Sequence 708, App
30	263	47.6	371	12	US-10-294-025-708	Sequence 708, App
31	263	47.6	371	14	US-10-012-896-708	Sequence 708, App
32	263	47.6	400	9	US-09-759-143-852	Sequence 852, App
33	263	47.6	400	9	US-09-780-669-852	Sequence 852, App
34	263	47.6	400	9	US-09-822-827-852	Sequence 852, App
35	263	47.6	400	10	US-09-895-793-852	Sequence 852, App
36	263	47.6	400	10	US-09-895-814-852	Sequence 852, App
37	263	47.6	400	12	US-10-144-678A-852	Sequence 852, App
38	263	47.6	400	12	US-10-294-025-852	Sequence 852, App
39	263	47.6	400	14	US-10-012-896-852	Sequence 852, App
40	238	43.0	246	12	US-10-144-678A-1028	Sequence 1028, Ap
41	238	43.0	246	12	US-10-294-025-1028	Sequence 1028, Ap
42	187	33.8	326	12	US-10-005-907-5	Sequence 5, Appli
43	187	33.8	355	12	US-10-144-678A-1011	Sequence 1011, Ap
44	187	33.8	355	12	US-10-294-025-1011	Sequence 1011, Ap
45	187	33.8	355	14	US-10-012-896-1011	Sequence 1011, Ap

ALIGNMENTS

RESULT 1

US-09-745-288-101

; Sequence 101, Application US/09745288

; Patent No. US20010018058A1

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE

; FILE REFERENCE: 210121.446d1

; CURRENT APPLICATION NUMBER: US/09/745,288

; CURRENT FILING DATE: 2000-12-19

; NUMBER OF SEQ ID NOS: 101

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 101

; LENGTH: 553

; TYPE: PRT

; ORGANISM: Homo sapien

US-09-745-288-101

Query Match	100.0%	Score 553;	DB 9;	Length 553;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 553;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	MYQRLWVSRLLRHRKAQLLNNLLTFGLVCLAAAGTTVPPLLLLEGVVEEKEMTWLIGT	60	
Db	1	MYQRLWVSRLLRHRKAQLLNNLLTFGLVCLAAAGTTVPPLLLLEGVVEEKEMTWLIGT	60	
Qy	61	PVLGLVCPVLLGSASDHWRGRRPFTWALSGLLSLFLIPRAGMLAGLCCPDRL	120	
Db	61	PVLGLVCPVLLGSASDHWRGRRPFTWALSGLLSLFLIPRAGMLAGLCCPDRL	120	
Qy	121	ELALLILGVLDFCGQVCFPLEALLSDLRDPDHCQAYSVYAFMISLGLCYLLPA	180	
Db	121	ELALLILGVLDFCGQVCFPLEALLSDLRDPDHCQAYSVYAFMISLGLCYLLPA	180	
Qy	181	IDWDTSAAPYLGTQECFLGLTLTIFLTCVAATLLVAEEAALGFTPEAGLSATSLSPH	240	

Db 181 IDWDTSAALPYLGTQBECLFGLLTLLIFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
QY 241 CCPCRLAFRNIGALLPRHLQCCMPRTLRRLFVAELCSMMALMTFTLFYDFVGBGL 300
Db 241 CCPCRLAFRNIGALLPRHLQCCMPRTLRRLFVAELCSMMALMTFTLFYDFVGBGL 300
QY 301 YQGVPAEPTGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPAEPTGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
QY 361 APFVAAGATCLSHSVAVVTASAAITGFTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAAITGFTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 2

US-09-838-785-2
; Sequence 2, Application US/09838785
; Patent No. US2002009455A1
; GENERAL INFORMATION:
; APPLICANT: Lau, Ted
; APPLICANT: Lin, Rick
; APPLICANT: Parkes, Debbie
; APPLICANT: Party, Gordon
; APPLICANT: Schneider, Douglas
; APPLICANT: Steinbrecher, Renate
; APPLICANT: Van Heuit, Pam T
; APPLICANT: Wu, John
; TITLE OF INVENTION: DNA Encoding a No. US2002009455A1el PROST 03
; FILE REFERENCE: 51831AUSM1
; CURRENT APPLICATION NUMBER: US/09/838,785
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/200,065
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-838-785-2

Query Match 100.0%; Score 553; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVQRLVWSRLRRKRAQLLLVNTLTFGLVCLAAAGITYVPPLLLEVGVEEKFMTWVLGIG 60
Db 1 MVQRLVWSRLRRKRAQLLLVNTLTFGLVCLAAAGITYVPPLLLEVGVEEKFMTWVLGIG 60
QY 61 PVGLVVCVPLLGASDHWGRYGRRRPFTWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVVCVPLLGASDHWGRYGRRRPFTWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILVGLLDFCCQVCTPLEALLSFLFRPDHCRQAYSVYAFWISLGGCLGYLLPA 180
Db 121 ELALLILVGLLDFCCQVCTPLEALLSFLFRPDHCRQAYSVYAFWISLGGCLGYLLPA 180
QY 181 IDWDTSAALPYLGTQBECLFGLLTLLIFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240

Db 181 IDWDTSAALPYLGTQBECLFGLLTLLIFLTCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
QY 241 CCPCRLAFRNIGALLPRHLQCCMPRTLRRLFVAELCSMMALMTFTLFYDFVGBGL 300
Db 241 CCPCRLAFRNIGALLPRHLQCCMPRTLRRLFVAELCSMMALMTFTLFYDFVGBGL 300
QY 301 YQGVPAEPTGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPAEPTGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
QY 361 APFVAAGATCLSHSVAVVTASAAITGFTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAAITGFTFSAIQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 3

US-09-759-143-113
; Sequence 113, Application US/09759143
; Patent No. US200202248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-759-143-113

Query Match 100.0%; Score 553; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVQRLVWSRLRRKRAQLLLVNTLTFGLVCLAAAGITYVPPLLLEVGVEEKFMTWVLGIG 60
Db 1 MVQRLVWSRLRRKRAQLLLVNTLTFGLVCLAAAGITYVPPLLLEVGVEEKFMTWVLGIG 60
QY 61 PVGLVVCVPLLGASDHWGRYGRRRPFTWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVVCVPLLGASDHWGRYGRRRPFTWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLILGVGLDPCGVCFTPLLEALLSDFRDPDHCRQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDPCGVCFTPLLEALLSDFRDPDHCRQAYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTSLAPYLGTQECFLGGLTLIFLTCVAATLLVAEEAALGPTTPEAGLSAPLSPH 240
DB 181 IDWDTSLAPYLGTQECFLGGLTLIFLTCVAATLLVAEEAALGPTTPEAGLSAPLSPH 240
QY 241 CCPCRLAFRNIGALLPRHLQCCMRPTLRLFLVLCVVAATLLVAEEAALGPTTPEAGLSAPLSPH 300
DB 241 CCPCRLAFRNIGALLPRHLQCCMRPTLRLFLVLCVVAATLLVAEEAALGPTTPEAGLSAPLSPH 300
QY 301 YQGVPRAEPTGTEARRHYDEGVMSGLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
DB 301 YQGVPRAEPTGTEARRHYDEGVMSGLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
QY 361 APVAAAGATCLSHSVAVVTASALTGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APVAAAGATCLSHSVAVVTASALTGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
QY 541 VVFDKSLAKYSA 553
DB 541 VVFDKSLAKYSA 553

RESULT 4

US-09-780-669-113
; Sequence 113, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien

US-09-780-669-113

Query Match 100.0%; Score 553; DB 9; Length 553;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVORLWVSLRLRRKKAQLLVNLLTFGLVCLAAAGITYVPPLLLEVGVBEKFTMWLGIG 60
DB 1 MVORLWVSLRLRRKKAQLLVNLLTFGLVCLAAAGITYVPPLLLEVGVBEKFTMWLGIG 60
QY 61 PVLGLVCVPLLSASDHWGRYVRRRPFIFWALSGLTLLSLFLIPRAGWLAGLCPDPRPL 120
DB 61 PVLGLVCVPLLSASDHWGRYVRRRPFIFWALSGLTLLSLFLIPRAGWLAGLCPDPRPL 120
QY 121 ELALLILGVGLDPCGVCFTPLLEALLSDFRDPDHCRQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDPCGVCFTPLLEALLSDFRDPDHCRQAYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTSLAPYLGTQECFLGGLTLIFLTCVAATLLVAEEAALGPTTPEAGLSAPLSPH 240
DB 181 IDWDTSLAPYLGTQECFLGGLTLIFLTCVAATLLVAEEAALGPTTPEAGLSAPLSPH 240
QY 241 CCPCRLAFRNIGALLPRHLQCCMRPTLRLFLVLCVVAATLLVAEEAALGPTTPEAGLSAPLSPH 300
DB 241 CCPCRLAFRNIGALLPRHLQCCMRPTLRLFLVLCVVAATLLVAEEAALGPTTPEAGLSAPLSPH 300
QY 301 YQGVPRAEPTGTEARRHYDEGVMSGLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
DB 301 YQGVPRAEPTGTEARRHYDEGVMSGLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
QY 361 APVAAAGATCLSHSVAVVTASALTGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APVAAAGATCLSHSVAVVTASALTGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
QY 541 VVFDKSLAKYSA 553
DB 541 VVFDKSLAKYSA 553

RESULT 5

US-09-030-606-113
; Sequence 113, Application US/09030606
; Patent No. US20020081580A1
; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF PROSTATE CANCER AND METHODS FOR
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,606
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.428C3

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-09-030-606-113

Query Match 100.0%; Score 553; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVVEEKFTMWLGIG 60
Db 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVVEEKFTMWLGIG 60

QY 61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLILGVGLDFCGQVCFPLEALLSDFLRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGQVCFPLEALLSDFLRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTSAALPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWDTSAALPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY 241 CCPCRLARAFNLGALLPRLHQLCCMRPTRLRLFVAELCSNMALMTFTLFTYDFVGEGL 300
Db 241 CCPCRLARAFNLGALLPRLHQLCCMRPTRLRLFVAELCSNMALMTFTLFTYDFVGEGL 300

QY 301 YQGVPAEPTGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPAEPTGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360

QY 361 APVPAAGATCLSHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVPAAGATCLSHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540

QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 6
US-09-822-827-113
; Sequence 113, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822, 827
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113

Query Match 100.0%; Score 553; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVVEEKFTMWLGIG 60
Db 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVVEEKFTMWLGIG 60

QY 61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLILGVGLDFCGQVCFPLEALLSDFLRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGQVCFPLEALLSDFLRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTSAALPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWDTSAALPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY 241 CCPCRLARAFNLGALLPRLHQLCCMRPTRLRLFVAELCSNMALMTFTLFTYDFVGEGL 300
Db 241 CCPCRLARAFNLGALLPRLHQLCCMRPTRLRLFVAELCSNMALMTFTLFTYDFVGEGL 300

QY 301 YQGVPAEPTGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPAEPTGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360

QY 361 APVPAAGATCLSHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVPAAGATCLSHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540

QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 7
US-09-115-453-113
; Sequence 113, Application US/09115453B
; Patent No. US20020090372A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND
; FILE REFERENCE: 210121.427C4
; CURRENT APPLICATION NUMBER: US/09/115, 453B
; CURRENT FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-115-453-113

Query Match 100.0%; Score 553; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
QY 1 MVQRLWSRLRRKKAQLLVNLITFGLEVCLAAGITYVPPLLELVGVEEKFTMWLGIG 60
|
|
|
Db 1 MVQRLWSRLRRKKAQLLVNLITFGLEVCLAAGITYVPPLLELVGVEEKFTMWLGIG 60
|
|
|
QY 61 PVLGLVCVPLGSGASDHWGRGRRRPFIIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
|
|
|
Db 61 PVLGLVCVPLGSGASDHWGRGRRRPFIIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
|
|
|
QY 121 ELALLIUGVLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISIGGCLGYLLPA 180
|
|
|
Db 121 ELALLIUGVLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISIGGCLGYLLPA 180
|
|
|
QY 181 IDWTSALAPYLGTQBECLFGLLTIFLTCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240
|
|
|
Db 181 IDWTSALAPYLGTQBECLFGLLTIFLTCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240
|
|
|
QY 241 CCPCRLARLAFRNIGALLPRHQLCCRMPTLRRLFVAELCSWMALMTFTLYTDFVGEGL 300
|
|
|
Db 241 CCPCRLARLAFRNIGALLPRHQLCCRMPTLRRLFVAELCSWMALMTFTLYTDFVGEGL 300
|
|
|
QY 301 YQGVPRABPTGTEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVQREGTRAVYLASVA 360
|
|
|
Db 301 YQGVPRABPTGTEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVQREGTRAVYLASVA 360
|
|
|
QY 361 AFPVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
|
|
|
Db 361 AFPVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
|
|
|
QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
|
|
|
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
|
|
|
QY 481 RVVPRGICLDLAILDSAPFLISOVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIYFATQ 540
|
|
|
Db 481 RVVPRGICLDLAILDSAPFLISOVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIYFATQ 540
|
|
|
QY 541 VVPDKSGLAKYSA 553
|
|
|
Db 541 VVPDKSGLAKYSA 553
|
|
|
RESULT 8
US-09-232-880-113
; Sequence 113, Application US/09232880
; Publication No. US20020182596A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF
; TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.428C6
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-232-880-113

Query Match 100.0%; Score 553; DB 10; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWSRLRRKKAQLLVNLITFGLEVCLAAGITYVPPLLELVGVEEKFTMWLGIG 60
|
|
|
Db 1 MVQRLWSRLRRKKAQLLVNLITFGLEVCLAAGITYVPPLLELVGVEEKFTMWLGIG 60
|
|
|
QY 61 PVLGLVCVPLGSGASDHWGRGRRRPFIIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
|
|
|
```

```
Db 61 PVLGLVCVPLGSGASDHWGRGRRRPFIIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
|
|
|
QY 121 ELALLIUGVLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISIGGCLGYLLPA 180
|
|
|
Db 121 ELALLIUGVLLDFCGQVCFPLEALLSDFRDPDHCROAYSVYAFMISIGGCLGYLLPA 180
|
|
|
QY 181 IDWTSALAPYLGTQBECLFGLLTIFLTCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240
|
|
|
Db 181 IDWTSALAPYLGTQBECLFGLLTIFLTCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240
|
|
|
QY 241 CCPCRLARLAFRNIGALLPRHQLCCRMPTLRRLFVAELCSWMALMTFTLYTDFVGEGL 300
|
|
|
Db 241 CCPCRLARLAFRNIGALLPRHQLCCRMPTLRRLFVAELCSWMALMTFTLYTDFVGEGL 300
|
|
|
QY 301 YQGVPRABPTGTEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVQREGTRAVYLASVA 360
|
|
|
Db 301 YQGVPRABPTGTEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVQREGTRAVYLASVA 360
|
|
|
QY 361 AFPVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
|
|
|
Db 361 AFPVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
|
|
|
QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
|
|
|
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
|
|
|
QY 481 RVVPRGICLDLAILDSAPFLISOVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIYFATQ 540
|
|
|
Db 481 RVVPRGICLDLAILDSAPFLISOVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIYFATQ 540
|
|
|
QY 541 VVPDKSGLAKYSA 553
|
|
|
Db 541 VVPDKSGLAKYSA 553
|
|
|
RESULT 9
US-09-895-793-113
; Sequence 113, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuxiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darlick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
```

US-09-895-793-113

Query Match 100.0%; Score 553; DB 10; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLVNLITFGLVCLAAAGITYVPPILLLEVGVEEKFTMWLGIG 60
DB 1 MVQRLWVSRLLRHRKAQLLVNLITFGLVCLAAAGITYVPPILLLEVGVEEKFTMWLGIG 60

QY 61 PVLGLVCVPLLSASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
DB 61 PVLGLVCVPLLSASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTLSALAPYLGTQBECLFGLLTILFVLCVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240
DB 181 IDWDTLSALAPYLGTQBECLFGLLTILFVLCVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240

QY 241 CCPCRLARLAFNRLGALLPRLHQLCCMRPTRLRLFLVAELCSMMALMTFTLFTYDFVGEGL 300
DB 241 CCPCRLARLAFNRLGALLPRLHQLCCMRPTRLRLFLVAELCSMMALMTFTLFTYDFVGEGL 300

QY 301 YQGVPRAPGTEARRHYDEGVGMGSLGLFLOCAISLVFSLVMDRLVQRTGTRAVYLASVA 360
DB 301 YQGVPRAPGTEARRHYDEGVGMGSLGLFLOCAISLVFSLVMDRLVQRTGTRAVYLASVA 360

QY 361 APVPAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APVPAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVPGRGICLDLAILDAPLLSOVAPSLFMGSIVOLQSQTAVMYSAAGLGLVYIFATQ 540
DB 481 RVVPGRGICLDLAILDAPLLSOVAPSLFMGSIVOLQSQTAVMYSAAGLGLVYIFATQ 540

QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 10

US-09-895-814-113
; Sequence 113, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Xuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-895-814-113

Query Match 100.0%; Score 553; DB 10; Length 553;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLVNLITFGLVCLAAAGITYVPPILLLEVGVEEKFTMWLGIG 60
DB 1 MVQRLWVSRLLRHRKAQLLVNLITFGLVCLAAAGITYVPPILLLEVGVEEKFTMWLGIG 60

QY 61 PVLGLVCVPLLSASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
DB 61 PVLGLVCVPLLSASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCROAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWDTLSALAPYLGTQBECLFGLLTILFVLCVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240
DB 181 IDWDTLSALAPYLGTQBECLFGLLTILFVLCVAATLLVAEEAALGPTPEAEGLSAPLSLSPH 240

QY 241 CCPCRLARLAFNRLGALLPRLHQLCCMRPTRLRLFLVAELCSMMALMTFTLFTYDFVGEGL 300
DB 241 CCPCRLARLAFNRLGALLPRLHQLCCMRPTRLRLFLVAELCSMMALMTFTLFTYDFVGEGL 300

QY 301 YQGVPRAPGTEARRHYDEGVGMGSLGLFLOCAISLVFSLVMDRLVQRTGTRAVYLASVA 360
DB 301 YQGVPRAPGTEARRHYDEGVGMGSLGLFLOCAISLVFSLVMDRLVQRTGTRAVYLASVA 360

QY 361 APVPAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APVPAAGATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVPGRGICLDLAILDAPLLSOVAPSLFMGSIVOLQSQTAVMYSAAGLGLVYIFATQ 540
DB 481 RVVPGRGICLDLAILDAPLLSOVAPSLFMGSIVOLQSQTAVMYSAAGLGLVYIFATQ 540

QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 11

US-10-144-678A-113
; Sequence 113, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Xuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.

APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Hepler, William T.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals y de Basols, Carlota
APPLICANT: Foy, Teresa M.
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Deng, Ta
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C28
CURRENT APPLICATION NUMBER: US/10/144, 678A
CURRENT FILING DATE: 2002-08-12
NUMBER OF SEQ ID NOS: 1033
SOFTWARE: Fast-Seq for Windows Version 3.0
SEQ ID NO 113
LENGTH: 553
TYPE: PRT
ORGANISM: Homo sapiens
US-10-144-678A-113

Query Match 100.0%; Score 553; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVORLWVSRLRRKKAQLLAVNLLTFGLEVCCLAAGTYVPPLLLVGVVEEKMVMVLGIG 60
DB 1 MVORLWVSRLRRKKAQLLAVNLLTFGLEVCCLAAGTYVPPLLLVGVVEEKMVMVLGIG 60
QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFTWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFTWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLDFCGQVCFPLEALLSDFRDPDHCROQYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFPLEALLSDFRDPDHCROQYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTSALAPYLGTQECFLGLLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
DB 181 IDWDTSALAPYLGTQECFLGLLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
QY 241 CCPCARLAFRNLGALLPRLHQLCCRMPTLRRLFVAELCSMWALMTFTLFTYDFVGEGL 300
DB 241 CCPCARLAFRNLGALLPRLHQLCCRMPTLRRLFVAELCSMWALMTFTLFTYDFVGEGL 300
QY 301 YQGVPRABPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
DB 301 YQGVPRABPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
QY 361 APVVAAGATCLSHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APVVAAGATCLSHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMYSAAGLGLVAIFYFATQ 540
DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMYSAAGLGLVAIFYFATQ 540
QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 13
US-10-294-025-113
Sequence 113, Application US/10294025
Publication No. US20030185830A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun

RESULT 12
US-10-005-907-13
Sequence 13, Application US/10005907
Publication No. US20030166881A1
GENERAL INFORMATION:
APPLICANT: Union Chimique Belge, S.A.
APPLICANT: No. US20030166881A1ka, Karl
APPLICANT: Pirozzi, Gregory
APPLICANT: Binstein, Richard
TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAST CELL
TITLE OF INVENTION: ACTIVATION
FILE REFERENCE: 053529-5005
CURRENT APPLICATION NUMBER: US/10/005,907
CURRENT FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
LENGTH: 553
TYPE: PRT
ORGANISM: Homo sapiens
US-10-005-907-13

Query Match 100.0%; Score 553; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVORLWVSRLRRKKAQLLAVNLLTFGLEVCCLAAGTYVPPLLLVGVVEEKMVMVLGIG 60
DB 1 MVORLWVSRLRRKKAQLLAVNLLTFGLEVCCLAAGTYVPPLLLVGVVEEKMVMVLGIG 60
QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFTWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFTWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLDFCGQVCFPLEALLSDFRDPDHCROQYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFPLEALLSDFRDPDHCROQYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTSALAPYLGTQECFLGLLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
DB 181 IDWDTSALAPYLGTQECFLGLLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
QY 241 CCPCARLAFRNLGALLPRLHQLCCRMPTLRRLFVAELCSMWALMTFTLFTYDFVGEGL 300
DB 241 CCPCARLAFRNLGALLPRLHQLCCRMPTLRRLFVAELCSMWALMTFTLFTYDFVGEGL 300
QY 301 YQGVPRABPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
DB 301 YQGVPRABPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFQTRAVYLASVA 360
QY 361 APVVAAGATCLSHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APVVAAGATCLSHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMYSAAGLGLVAIFYFATQ 540
DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMYSAAGLGLVAIFYFATQ 540
QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 13
US-10-294-025-113
Sequence 113, Application US/10294025
Publication No. US20030185830A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun

; APPLICANT: Stolk, John A.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294,025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-025-113

Query Match 100.0%; Score 553; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVORLWVSLRHRKAQLLVNLTGFEVCLAAAGTYVPPLLLEVGVEEKMVMVLGIG 60
DB 1 MVORLWVSLRHRKAQLLVNLTGFEVCLAAAGTYVPPLLLEVGVEEKMVMVLGIG 60
QY 61 PVGLVVCVPLGASDHWGRYGRRRPFIWALSGLISLFLIPRAGWLAGLLCPDPRPL 120
DB 61 PVGLVVCVPLGASDHWGRYGRRRPFIWALSGLISLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTSALAPYLGTQBECLFGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
DB 181 IDWDTSALAPYLGTQBECLFGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
QY 241 CCPCARLAFRNLGALLPRHLQCCMRPTLRLLFVAELCSWMLMTFTLFTYDFVGEGL 300
DB 241 CCPCARLAFRNLGALLPRHLQCCMRPTLRLLFVAELCSWMLMTFTLFTYDFVGEGL 300
QY 301 YQGVPAEPGTEARRHYDEGVRMGSIGLFLQCAISLVFSLVMDRLVQRFTRAVYLASVA 360
DB 301 YQGVPAEPGTEARRHYDEGVRMGSIGLFLQCAISLVFSLVMDRLVQRFTRAVYLASVA 360
QY 361 APPVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APPVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAYIFATQ 540
DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAYIFATQ 540
QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 14

US-10-012-896-113
; Sequence 113, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-113

Query Match 100.0%; Score 553; DB 14; Length 553;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVORLWVSLRHRKAQLLVNLTGFEVCLAAAGTYVPPLLLEVGVEEKMVMVLGIG 60
DB 1 MVORLWVSLRHRKAQLLVNLTGFEVCLAAAGTYVPPLLLEVGVEEKMVMVLGIG 60
QY 61 PVGLVVCVPLGASDHWGRYGRRRPFIWALSGLISLFLIPRAGWLAGLLCPDPRPL 120
DB 61 PVGLVVCVPLGASDHWGRYGRRRPFIWALSGLISLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWDTSALAPYLGTQBECLFGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
DB 181 IDWDTSALAPYLGTQBECLFGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
QY 241 CCPCARLAFRNLGALLPRHLQCCMRPTLRLLFVAELCSWMLMTFTLFTYDFVGEGL 300
DB 241 CCPCARLAFRNLGALLPRHLQCCMRPTLRLLFVAELCSWMLMTFTLFTYDFVGEGL 300
QY 301 YQGVPAEPGTEARRHYDEGVRMGSIGLFLQCAISLVFSLVMDRLVQRFTRAVYLASVA 360
DB 301 YQGVPAEPGTEARRHYDEGVRMGSIGLFLQCAISLVFSLVMDRLVQRFTRAVYLASVA 360
QY 361 APPVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APPVAAGATCLSHSVAVVTASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAYIFATQ 540
DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAYIFATQ 540
QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 15

US-10-010-940-113
; Sequence 113, Application US/10010940
; Publication No. US2003008062A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427D3
; CURRENT APPLICATION NUMBER: US/10/010,940
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien

US-10-010-940-113

Query Match 100.0%; Score 553; DB 15; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MVQRLWVSRLRRKQALLVNLITFGLVCLAAAGITYPPLLLEVGVEEKFTMTVLIG 60
Db	1	MVQRLWVSRLRRKQALLVNLITFGLVCLAAAGITYPPLLLEVGVEEKFTMTVLIG 60
Qy	61	PVLGLVCVPLGSGASDHWGRYGRRRPFIWALSIGLILSLFLIPRAGWLAGLLCPDPRPL 120
Db	61	PVLGLVCVPLGSGASDHWGRYGRRRPFIWALSIGLILSLFLIPRAGWLAGLLCPDPRPL 120
Qy	121	ELALLILGVGLLFCGQVCFPTLEALLSDFRDPDHCRCQAYSYYAFWISLGGCLGYLLPA 180
Db	121	ELALLILGVGLLFCGQVCFPTLEALLSDFRDPDHCRCQAYSYYAFWISLGGCLGYLLPA 180
Qy	181	IDWDTGALAPYLGTQBECLFGLITLIFLTCVAATLLVABEAAALGPTPEAGLSAPLSPH 240
Db	181	IDWDTGALAPYLGTQBECLFGLITLIFLTCVAATLLVABEAAALGPTPEAGLSAPLSPH 240
Qy	241	CCPCRLAFRNIGALLPRHLQCCRMPTRELFLVAELCSWMLMTFTLFYTDVFVGEGL 300
Db	241	CCPCRLAFRNIGALLPRHLQCCRMPTRELFLVAELCSWMLMTFTLFYTDVFVGEGL 300
Qy	301	YQGVPRAEPTGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFCTRAVYLASVA 360
Db	301	YQGVPRAEPTGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFCTRAVYLASVA 360
Qy	361	APFVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLTASLYHREKQVFLPKYRGDTGG 420
Db	361	APFVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLTASLYHREKQVFLPKYRGDTGG 420
Qy	421	ASSEDLSMTSFLPGPKGAPFPNGHVGAGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db	421	ASSEDLSMTSFLPGPKGAPFPNGHVGAGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Qy	481	RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMWSAAGLGLVAYIFATQ 540
Db	481	RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMWSAAGLGLVAYIFATQ 540
Qy	541	VVPDKSDLAKYSA 553
Db	541	VVPDKSDLAKYSA 553

Search completed: December 3, 2003, 17:25:48
Job time : 37 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore Version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 17:20:47 ; Search time 40 Seconds
(without alignments)
3567.577 Million cell updates/sec

Title: US-09-593-793A-113

Perfect score: 553

Sequence: 1 MVQRLWVSRLLRHRKAQLL.....AIYFATQVVPDKSLAKYSA 553

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_23.*

- 1: sp_archea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	553	100.0	553	4 Q96JT2	Q96jt2 homo sapien
2	158	28.6	553	6 Q95KI5	Q95ki5 macaca fasc
3	131	23.7	501	6 Q95XC5	Q95kc5 macaca fasc
4	59	10.7	450	11 Q8K252	Q8k252 mus musculus
5	59	10.7	553	11 Q8KH07	Q8kh07 mus musculus
6	33	6.0	112	11 Q8RL10	Q8rl10 mus musculus
7	10	1.8	520	4 Q8NPF8	Q8nfh8 homo sapien
8	10	1.8	521	4 Q43428	Q43428 homo sapien
9	10	1.8	659	4 Q8NFT5	Q8nf15 homo sapien
10	9	1.6	305	16 Q8DL01	Q8dl01 synechococc
11	9	1.6	488	10 Q9FV92	Q9fv92 solanum tub
12	9	1.6	500	10 Q9FVJ6	Q9fvj6 lycopersico
13	9	1.6	631	16 Q8E924	Q8e924 shewanella
14	8	1.4	68	15 Q9QJC6	Q9qjc6 human immun
15	8	1.4	115	15 Q90UA8	Q90ua8 human immun
16	8	1.4	120	2 Q93PB7	Q93pb7 microscilla

17	8	1.4	135	10 Q9LTS6	Q9lts6 arabidopsis
18	8	1.4	135	10 Q8LDI6	Q8ldi6 arabidopsis
19	8	1.4	142	16 Q8VVF9	Q8vvf9 anabaena sp
20	8	1.4	176	5 Q94075	Q94075 caenorhabdi
21	8	1.4	194	10 Q9LI23	Q9liz3 oryza sativ
22	8	1.4	232	16 Q8R8K3	Q8r8k3 thermoanaer
23	8	1.4	234	17 Q8TX99	Q8txy9 methanopyru
24	8	1.4	242	2 Q9XBN0	Q9xbn0 rhodococcus
25	8	1.4	260	2 Q54386	Q54386 streptomyce
26	8	1.4	282	16 Q8DBF6	Q8dbf6 vibrio vuln
27	8	1.4	315	16 Q98GI0	Q98gi0 rhizobium l
28	8	1.4	354	2 Q59257	Q59257 bacillus sp
29	8	1.4	378	10 Q942N6	Q942n6 oryza sativ
30	8	1.4	404	10 Q9C9D2	Q9c9d2 arabidopsis
31	8	1.4	405	16 Q8E3Z0	Q8e3z0 streptococ
32	8	1.4	405	16 Q8DYC6	Q8dyc6 streptococ
33	8	1.4	412	16 Q9A686	Q9a686 caulobacter
34	8	1.4	445	10 Q9C6A1	Q9c6a1 arabidopsis
35	8	1.4	457	2 Q9F499	Q9f499 erwinia amy
36	8	1.4	474	10 Q04516	Q04516 arabidopsis
37	8	1.4	480	16 Q05438	Q05438 mycobacteri
38	8	1.4	498	16 Q9KZ81	Q9kz81 streptomyce
39	8	1.4	501	10 Q85929	Q85929 daucus caro
40	8	1.4	510	10 Q9FE59	Q9fe59 arabidopsis
41	8	1.4	510	10 Q9M3R4	Q9m3r4 arabidopsis
42	8	1.4	594	10 Q80605	Q80605 arabidopsis
43	8	1.4	596	16 Q9A2N0	Q9a2n0 caulobacter
44	8	1.4	604	10 Q9FVL6	Q9fv16 lycopersico
45	8	1.4	607	10 Q8LPM4	Q8lpm4 citrus sine

ALIGNMENTS

RESULT 1

Q96JT2 ID Q96JT2 PRELIMINARY; PRT; 553 AA.

AC Q96JT2; Q96JT2 (TREMREL. 19, Created)

DT 01-DEC-2001 (TREMREL. 19, Last sequence update)

DT 01-DEC-2001 (TREMREL. 19, Last annotation update)

DE Prostelin.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Prostate;

RX MEDLINE=21139094; PubMed=11245466;

RA Xu J., Kalos M., Stolk J.A., Zasloff E.J., Zhang X., Houghton R.L.,

RA Filho A.M., Nolasco M., Badaro R., Reed S.G.;

RT "Identification and characterization of prostate, a novel prostate-

RT specific protein.";

RL Cancer Res. 61:1563-1568(2001).

DR EMBL, AY033593; AAK54386.1; -.

SQ SEQUENCE 553 AA; 59322 MW; 0AFA23FEC742A667 CRC64;

Query Match 100.0%; Score 553; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLLRHRKAQLLNLTPGLEVCIAAGITVPPPLLELVGVGEKFTMWLGIG 60

DB 1 MVQRLWVSRLLRHRKAQLLNLTPGLEVCIAAGITVPPPLLELVGVGEKFTMWLGIG 60

QY 61 PVGLGVCVPLLGASDHWGRYGRRRPFTWALSIGLLSLFLIPRAGWLAGLLCPDPRPL 120

DB 61 PVGLGVCVPLLGASDHWGRYGRRRPFTWALSIGLLSLFLIPRAGWLAGLLCPDPRPL 120

QY 121 ELALITLGVGLDFGCGQVCFPLEALLSLDFRDPDHCRCQAYSFYAFMISLGCLGYLLPA 180

DB 121 ELALITLGVGLDFGCGQVCFPLEALLSLDFRDPDHCRCQAYSFYAFMISLGCLGYLLPA 180

```
QY 181 IDWDTSAALPYLCTOBECLFGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLPH 240
Db 181 IDWDTSAALPYLCTOBECLFGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPLSLPH 240
QY 241 CCPCRLAFRNLCALPRHLQCCMRPRLRLFLVAELCSWMLMTFTLFYDFVGEGL 300
Db 241 CCPCRLAFRNLCALPRHLQCCMRPRLRLFLVAELCSWMLMTFTLFYDFVGEGL 300
QY 301 YQGVPAEPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPAEPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
QY 361 AFVAAGATCLSHVAVTASALTGTFTSALQILPYTLASLYHREKQVLPKYRGDTGG 420
Db 361 AFVAAGATCLSHVAVTASALTGTFTSALQILPYTLASLYHREKQVLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPTPEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPTPEA 480
QY 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
Db 481 RVVPGRGICLDLAILDSAPLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 2
Q95KI5
ID Q95KI5 PRELIMINARY; PRT; 553 AA.
AC Q95KI5;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DE Hypothetical 59.4 kDa protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Temporal cortex;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB060851; BAB46871.1; -
KW Hypothetical protein.
SQ SEQUENCE 553 AA; 59392 MW; 0718F3A91PB3BF1E CRC64;

Query Match 28.6%; Score 158; DB 6; Length 553;
Best Local Similarity 100.0%; Pred. No. 5.8e-144;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 VPLGSGASHWRCYGRRRPFIWALSIGLLSLFLIPRAGWLAGLCPDPRPLEALLIL 127
Db 68 VPLGSGASHWRCYGRRRPFIWALSIGLLSLFLIPRAGWLAGLCPDPRPLEALLIL 127
QY 128 GVGLLDFCGQVCTPTEALLSDLFRDPDHCQRAYSVAFMISLGGCIGYLLPAIDWDTSA 187
Db 128 GVGLLDFCGQVCTPTEALLSDLFRDPDHCQRAYSVAFMISLGGCIGYLLPAIDWDTSA 187
QY 188 LAPYLCTOBECLFGLTLTFLTCVAATLLVAEEAALGP 225
Db 188 LAPYLCTOBECLFGLTLTFLTCVAATLLVAEEAALGP 225

RESULT 3
```

```
Q95KC5
ID Q95KC5 PRELIMINARY; PRT; 501 AA.
AC Q95KC5;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DE Hypothetical 53.4 kDa protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Medulla oblongata;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062977; BAB60745.1; -
KW Hypothetical protein.
SQ SEQUENCE 501 AA; 53447 MW; 8C554BBD04EE0470 CRC64;

Query Match 23.7%; Score 131; DB 6; Length 501;
Best Local Similarity 100.0%; Pred. No. 7.6e-118;
Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 422 SSEDLSMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPTPEAR 481
Db 370 SSEDLSMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPTPEAR 429
QY 482 VVPGRGICLDLAILDSAPLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQV 541
Db 430 VVPGRGICLDLAILDSAPLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQV 489
QY 542 VFDKSDLAKYS 552
Db 490 VFDKSDLAKYS 500

RESULT 4
Q8K252
ID Q8K252 PRELIMINARY; PRT; 450 AA.
AC Q8K252;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DE Hypothetical protein (Fragment).
GN 2210413P12RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC034084; AAH34084.1; -
DR MGD; MGI:1922082; 2210413P12RIK.
DR InterPro; IPR005805; Rieseke.
DR PROSITE; PS00200; RIESEKE_2; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 450 AA; 48261 MW; 72C703AEC131302C CRC64;

Query Match 10.7%; Score 59; DB 11; Length 450;
Best Local Similarity 100.0%; Pred. No. 3.8e-48;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 QLCCRMPTRLRLFLVAELCSWMLMTFTLFYDFVGEGLYQGVPAEPGTEARRHYDEG 320
Db 159 QLCCRMPTRLRLFLVAELCSWMLMTFTLFYDFVGEGLYQGVPAEPGTEARRHYDEG 217
```



```

RESULT 5
Q8KOH7 ID Q8KOH7 PRELIMINARY; PRT; 553 AA.
AC Q8KOH7
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein (PROTEIN homolog).
GN 2210413P12RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Liver;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
PC SPRAIN-C57BL/6J; Tissue=Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; BC031381; AH31381.1; -
DR EMBL; AK035428; BAC29063.1; -
DR MGD; MGI:1922082; 2210413P12RIK.
DR InterPro; IPR005805; Rieske.
DR PROSITE; PS00200; RIESKE_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 553 AA; 59742 MW; GA474FEA630B4B13 CRC64;

Query Match 10.7%; Score 59; DB 11; Length 553;
Best Local Similarity 100.0%; Pred. No. 4.6e-48;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SLLLRHKAQLLVNLLTFGLVCLAGITVPPPLLEVGVEKFTMTVLGIGPVLGLV 66
Db 8 SLLLRHKAQLLVNLLTFGLVCLAGITVPPPLLEVGVEKFTMTVLGIGPVLGLV 66

RESULT 6
Q8RI10 ID Q8RI10 PRELIMINARY; PRT; 112 AA.
AC Q8RI10
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 11.4 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Colon;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024519; AAH24519.1; -
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 112 AA; 11393 MW; B9D9DE79FEF09PB7 CRC64;

Query Match 6.0%; Score 33; DB 11; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.7e-23;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 GRGICLDLAIDSAFLLSQVAPSLFPGSIIVQLS 517

```

```

Db 44 GRGICLDLAIDSAFLLSQVAPSLFPGSIIVQLS 76

RESULT 7
Q8NFH8 ID Q8NFH8 PRELIMINARY; PRT; 520 AA.
AC Q8NFH8
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RALBP1 associated Eps domain containing protein 2b.
GN REPS2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Oosterhoff J.K.; Penninkhof F.; Brinkmann A.O.; Grootegeod A.J.;
RA Blok L.J.;
RT "Role of REPS2/POB1 in progression of prostate cancer.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF512951; AAM43953.1; -
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000261; EPS15_homology.
DR Pfam; PF00036; ehand; 1.
DR SMART; SM00027; EH; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00031; EH; 1.
SQ SEQUENCE 520 AA; 57773 MW; C5B4F557D58A444D CRC64;

Query Match 1.8%; Score 10; DB 4; Length 520;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 451 SGLLPPPPAL 460
Db 369 SGLLPPPPAL 378

RESULT 8
Q43428 ID Q43428 PRELIMINARY; PRT; 521 AA.
AC Q43428
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RalBP1-interacting protein.
GN POB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98086250; PubMed=9422736;
RA Ikeda M.; Ishida O.; Hinoi T.; Kishida S.; Kikuchi A.;
RT "Identification and characterization of a novel protein interacting
RT with Ral-binding protein 1, a putative effector protein of Ral.";
RL J. Biol. Chem. 273:814-821 (1998).
DR EMBL; AF010233; AAC02901.1; -
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000261; EPS15_homology.
DR Pfam; PF00036; ehand; 1.
DR SMART; SM00027; EH; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00031; EH; 1.
SQ SEQUENCE 521 AA; 57901 MW; F5FCF51B56961026 CRC64;

Query Match 1.8%; Score 10; DB 4; Length 521;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 451 SGLPPPPAL 460
Db 370 SGLPPPPAL 379

RESULT 9
O8NFI5 PRELIMINARY; PRT; 659 AA.
ID Q8NFI5
AC Q8NFI5
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RALBP1 associated Eps domain containing protein 2a.
GN REPS2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Oosterhoff J.K., Penninkhof F., Brinkmann A.O., Grootegeed A.J.,
RA Blok L.J.;
RT "Role of REPS2/POB1 in progression of prostate cancer.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL; AF511533; AA43933.1; -.
DR Genew; HGNC:9963; REPS2.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand; 1.
DR SMART; SM00027; EH; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PSS0031; EH; 2.
SQ SEQUENCE 659 AA; 71405 MW; 527A658BDFB4582B CRC64;

Query Match 1.8%; Score 10; DB 4; Length 659;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 451 SGLPPPPAL 460
Db 508 SGLPPPPAL 517

RESULT 10
Q8DL01 PRELIMINARY; PRT; 305 AA.
ID Q8DL01
AC Q8DL01
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Indole-3-glycerol phosphate synthase.
GN TRPC OR TLL0698.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Saeemoto S.,
RA Watanabe A., Iriiguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR ENBL; AP005371; BAC08249.1; -.
KW Complete proteome.
SQ SEQUENCE 305 AA; 33953 MW; A6B394C1A3CE718B CRC64;

Query Match 1.6%; Score 9; DB 16; Length 305;
Best Local Similarity 100.0%; Pred. No. 7.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 VAAGATCLS 372
Db 107 VAAGATCLS 115

RESULT 11
Q9FV92 PRELIMINARY; PRT; 488 AA.
ID Q9FV92
AC Q9FV92
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Sucrose transporter SUT4 (Fragment).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20407148; PubMed=10948254;
RA Weise A., Barker L., Kuehn C., Lalonde S., Buschmann H., Frommer W.B.,
RA Ward J.M.;
RT "A new subfamily of sucrose transporters, SUT4, with low affinity/high
capacity localized in enucleate sieve elements of plants.";
RL Plant Cell 12:1345-1356(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR ENBL; AF237780; AAG25923.1; -.
DR InterPro; IPR005828; Sub transporter..
DR Pfam; PF00083; sugar_cr; 1.
DR TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 488
SQ SEQUENCE 488 AA; 52840 MW; 0800388914ADAF77 CRC64;

Query Match 1.6%; Score 9; DB 10; Length 488;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 GVRMGSLGL 328
Db 319 GVRMGSLGL 327

RESULT 12
Q9FVJ6 PRELIMINARY; PRT; 500 AA.
ID Q9FVJ6
AC Q9FVJ6
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Sucrose transporter.
OS SUT4.
OC Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Moneymaker;
RX MEDLINE=20407148; PubMed=10948254;
RA Weise A., Barker L., Kuehn C., Lalonde S., Buschmann H., Frommer W.B.,
RA Ward J.M.;
RT "A new subfamily of sucrose transporters, SUT4, with low affinity/high
capacity localized in enucleate sieve elements of plants.";
RL Plant Cell 12:1345-1355(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR ENBL; AF176950; AAG09270.1; -.

```

```
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005989; Suc/H_symport.
DR Pfam; PF00083; sugar tr; 1.
DR TIGRPFAMs; TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
SQ SEQUENCE 500 AA; 54274 MW; BDFPFE7468DA0110 CRC64;

Query Match 1.6%; Score 9; DB 10; Length 500;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 320 GVRMGSLGL 328
Db 326 GVRMGSLGL 334

RESULT 13
Q8E9Z4 PRELIMINARY; PRT; 631 AA.
AC Q8E9Z4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MSHA biogenesis protein MshH.
GN MSHH OR SO4116.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadales; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RA MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.P., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward M., Methe B., Clayton R.A.,
RA Deboy R., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA Madupu R., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Vamathevan J., Weidman J.D., Umayam L.A., White O., Wolf A.M.,
RA Mueller J., Khouri H., Gill J., Imprim M., Lee K., Berry K., Lee C.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neale K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015843; AAN57089.1; -.
DR TIGR; SO4116; -.
KW Complete proteome.
SQ SEQUENCE 631 AA; 72176 MW; 8BE730F9D77E4FB1 CRC64;

Query Match 1.6%; Score 9; DB 16; Length 631;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 124 LLILGVGLL 132
Db 130 LLILGVGLL 138

RESULT 14
Q9QJC6 PRELIMINARY; PRT; 68 AA.
AC Q9QJC6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope protein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=151N;

RX MEDLINE=20170331; PubMed=10708058;
RA Casado C., Urtaeun I., Martin-Walther M.V., Garcia S., Rodriguez C.,
RA del Romero J., Lopez-Galindez C.;
RT "Genetic analysis of HIV-1 samples from Spain.";
RL J. Acquir. Immune Defic. Syndr. 23:68-74(2000).
DR EMBL; AF152836; AAF08474.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 68
SQ SEQUENCE 68 AA; 7538 MW; 76B0F5F1734650C7 CRC64;

Query Match 1.4%; Score 8; DB 15; Length 68;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 513 IVQLSQSV 520
Db 3 IVQLSQSV 10

RESULT 15
Q9QUA8 PRELIMINARY; PRT; 115 AA.
AC Q9QUA8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21192328; PubMed=11294662;
RA Gunthard H.F., Havlik R., Fisch S., Zhang Z.-Q., Eron J.,
RA Mellors J., Gulick R., Frost S.D., Leigh Brown A.J., Schleif W.,
RA Valentine F., Jonas L., Meibohm A., Ignacio C.C., Isaacs R.,
RA Gamagami R., Emini E., Haase A., Richman D.D., Wong J.K.;
RT "Residual human immunodeficiency virus (HIV) type 1 RNA and DNA in
RT lymph nodes and HIV RNA in genital secretions and in cerebrospinal
RT fluid after suppression of viremia for 2 years.";
RL J. Infect. Dis. 183:1318-1327(2001).
DR EMBL; AF337292; AAKS6214.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 115
SQ SEQUENCE 115 AA; 12907 MW; 4490635198880152 CRC64;

Query Match 1.4%; Score 8; DB 15; Length 115;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 513 IVQLSQSV 520
Db 38 IVQLSQSV 45

Search completed: December 3, 2003, 17:24:25
Job time : 41 secs
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 17:21:22 ; Search time 21 seconds
(without alignments)
2532.442 Million cell updates/sec

Title: US-09-593-793A-113
Perfect score: 553
Sequence: 1 MVQRLWVSRLLRHRKAQLL.....AIYFATQVVFVDSKLAKYSA 553

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	1.6	1609	2 E87243	probable cation tr
2	9	1.6	1632	2 C70752	probable ctpi prot
3	8	1.4	137	2 S16222	bombinin H precurs
4	8	1.4	142	2 AE2058	hypothetical prote
5	8	1.4	176	2 T25087	hypothetical prote
6	8	1.4	223	2 JX0222	ubiquitin thiolest
7	8	1.4	260	1 JS0635	rRNA (adenine-N6-)
8	8	1.4	328	2 A48971	1S1380 polypeptide
9	8	1.4	354	1 S51779	endo-1,4-beta-xyla
10	8	1.4	412	2 G87522	hypothetical prote
11	8	1.4	445	2 C96769	unknown protein F9
12	8	1.4	455	1 WQEST	phosphotransferase
13	8	1.4	456	2 S62331	phosphotransferase
14	8	1.4	474	2 A86234	hypothetical prote
15	8	1.4	480	2 D70600	probable polynucle
16	8	1.4	501	2 T14339	sucrose-proton tra
17	8	1.4	594	2 G84441	probable sucrose-p
18	8	1.4	596	2 E87686	succinate dehydrog
19	8	1.4	787	2 S47245	ActA protein - Lis
20	8	1.4	1080	2 T43164	factA protein - Li
21	8	1.4	1804	2 T34518	nestin - golden ha
22	7	1.3	68	2 AD3215	hypothetical prote
23	7	1.3	83	2 A71471	hypothetical prote
24	7	1.3	87	2 C46449	hypothetical prote
25	7	1.3	89	1 CCS66	cytochrome c6 - Sp
26	7	1.3	102	2 E87453	hypothetical prote
27	7	1.3	122	2 T49358	related to glycine
28	7	1.3	128	2 G72807	gp64 protein - Myc
29	7	1.3	130	2 S74529	hypothetical prote

ALIGNMENTS

RESULT 1

E87243
probable cation transport ATPase ML2671 [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: E87243
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hori
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A;Title: Massive gene decay in the leprosy bacillus
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: E87243
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1609 <STO>
A;Cross-references: GB:AL450380; NID:G13093871; PIDN:CAC32203.1; GSPDB:GN00147
C;Genetics:
A;Gene: ML2671

Query Match 1.6%; Score 9; DB 2; Length 1609;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QLLLVNLLT 25
|||
Db 1425 QLLLVNLLT 1433

RESULT 2

C70752
probable ctpi protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: C70752
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: C70752
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1632 <COL>
A;Cross-references: GB:Z74410; GB:AL123456; NID:G3261600; PIDN:CAA98940.1; PID:el299890;
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: ctpi
C;Superfamily: ATPase nucleotide-binding domain homology

thymic shared anti
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
conserved hypothet
hypothetical prote
hypothetical prote
hypothetical prote
pupi protein - Pse
probable arok prot

30 7 1.3 136 2 I49013
31 7 1.3 141 2 C83768
32 7 1.3 146 2 E70845
33 7 1.3 149 1 B63061
34 7 1.3 149 2 B91085
35 7 1.3 149 2 D85930
36 7 1.3 149 2 AF0127
37 7 1.3 149 2 AF0861
38 7 1.3 151 1 S45108
39 7 1.3 151 2 A96493
40 7 1.3 154 2 H83160
41 7 1.3 155 2 S29280
42 7 1.3 160 2 S77324
43 7 1.3 173 2 F70481
44 7 1.3 173 2 S46355
45 7 1.3 176 2 G70658

F;1237-1406/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 1.6%; Score 9; DB 2; Length 1632;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QLLVNLTLT 25

DB 1430 QLLVNLTLT 1438

RESULT 3

S16222
bomblinin H precursor - yellow-bellied toad
N;Contains: Bomblinin H
C;Species: Bombina variegata (yellow-bellied toad)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jul-2000
C;Accession: S16222; A44581
R;Simmaco, M.; Barra, D.; Chiarini, F.; Noviello, L.; Melchiorri, P.; Kreil, G.; Richters, J. Biochem. 199, 217-222, 1991
A;Title: A family of bomblinin-related peptides from the skin of Bombina variegata.
A;Reference number: S16222; MUID:91293126; PMID:1712299
A;Accession: S16222
A;Molecule type: mRNA
A;Residues: 1-137 <STM>
A;Cross-references: GB:X59695; NID:962568; PIDN:CAM42216.1; PID:962569
R;Mignogna, G.; Simmaco, M.; Kreil, G.; Barra, D.
EMBO J. 12, 4829-4832, 1993
A;Title: Antibacterial and haemolytic peptides containing D-alloisoleucine from the skin of Bombina variegata.
A;Reference number: S39612; MUID:94038967; PMID:8223491
A;Accession: A44581
A;Status: preliminary
A;Molecule type: protein
A;Residues: 117-136 <MIG>
C;Superfamily: bomblinin H precursor
C;Keywords: amidated carboxyl end; D-amino acid
F;117-136/Product: bomblinin H #status experimental <MAT>
F;118/Modified site: D-allo-isoleucine (Ile) (in mature form) #status experimental
F;136/Modified site: amidated carboxyl end (Ile) (amide in mature form from following gl

Query Match 1.4%; Score 8; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 IGPVLGLV 66

DB 118 IGPVLGLV 125

RESULT 4

AE2058
hypothetical protein alr2019 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AE2058
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2058
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-142 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA073718.1; PID:gl7131109; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr2019

Query Match 1.4%; Score 8; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 LLSLFLIP 104
DB 86 LLSLFLIP 93

RESULT 5

T25087
hypothetical protein T21H8.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T25087
R;Gajadaty, S.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z19979
A;Accession: T25087
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-176 <WIL>
A;Cross-references: EMBL:Z78546; PIDN:CAB01770.1; GSPDB:GN00028; CESP:T21H8.5
A;Experimental source: clone T21H8
C;Genetics:
A;Gene: CESP:T21H8.5
A;Map position: X
A;Introns: 25/1; 83/1; 112/1
C;Superfamily: Caenorhabditis elegans hypothetical protein T21H8.5

Query Match 1.4%; Score 8; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 453 LLPPPPAL 460
DB 51 LLPPPPAL 58

RESULT 6

JX0222
ubiquitin thiolesterase (EC 3.1.2.15) PGP9.5 - rat
N;Alternate names: ubiquitin carboxy-terminal esterase; ubiquitin carboxyl-terminal hydrolase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 10-Dec-1999
C;Accession: JX0222
R;Kajimoto, Y.; Hashimoto, T.; Shirai, Y.; Nishino, N.; Kuno, T.; Tanaka, C.
J. Biochem. 112, 28-32, 1992
A;Title: cDNA cloning and tissue distribution of a rat ubiquitin carboxyl-terminal hydrolase
A;Reference number: JX0222; MUID:93054416; PMID:1331034
A;Accession: JX0222
A;Molecule type: mRNA
A;Residues: 1-223 <KAJ>
A;Cross-references: DDBJ:D01109
A;Experimental source: brain
C;Superfamily: human ubiquitin thiolesterase
C;Keywords: thiolester hydrolase

Query Match 1.4%; Score 8; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 GASSDLSL 427
DB 186 GASSDLSL 193

RESULT 7

JS0635
rRNA (adenine-N6-)-methyltransferase (EC 2.1.1.48) - Streptomyces lividans
C;Species: Streptomyces lividans
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-May-2000
C;Accession: JS0635
R;Jenkins, G.; Cundliffe, E.
Gene 108, 55-62, 1991
A;Title: Cloning and characterization of two genes from Streptomyces lividans that confer

A;Reference number: JS0635; MUID:92104506; PMID:1761231
A;Accession: JS0635
A;Molecule type: DNA
A;Residues: 1-260 <JRN>
A;Cross-references: GB:W74717; NID:gl53345; PIDN:AAA26779.1; PID:gl53346
A;Experimental source: strain TK21
C;Comment: This enzyme confers resistance to lincomycin.
C;Comment: This enzyme catalyzes the monomethylation of a specific adenosine within 23S
C;Genetics:
A;Gene: lrm
C;Superfamily: rRNA (adenine-N6-)-methyltransferase
C;Keywords: methyltransferase; S-adenosylmethionine

Query Match 1.4%; Score 8; DB 1; Length 260;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 PPLLEVG 47
|||||
DB 36 PPLLEVG 43

RESULT 8
A48971
IS1380 polypeptide homolog - Bradyrhizobium japonicum
C;Species: Bradyrhizobium japonicum
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995
C;Accession: A48971
R;Judd, A.K.; Sadowsky, M.J.
Appl. Environ. Microbiol. 59, 1656-1661, 1993
A;Title: The Bradyrhizobium japonicum serocenter 123 hyperreiterated DNA region, HRS1.
A;Reference number: A48971; MUID:93298005; PMID:8390818
A;Accession: A48971
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-328 <JUD>
A;Experimental source: USDA 424
A;Note: sequence extracted from NCBI backbone (NCBIN:133925, NCBIP:133926)

Query Match 1.4%; Score 8; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 GLSAPSL 238
|||||
DB 275 GLSAPSL 282

RESULT 9
S51779
endo-1,4-beta-xylanase (EC 3.2.1.8) Y precursor - Bacillus sp. (strain YA-335)
N;Alternate names: xylanase Y
C;Species: Bacillus sp.
A;Variety: strain YA-335
C;Date: 15-Jul-1995 #sequence_revision 22-Nov-1996 #text_change 16-Jun-2000
C;Accession: S51779; S48127
R;Hyun Ju, Y.
submitted to the EMBL Data Library, December 1990
A;Reference number: S51779
A;Accession: S51779
A;Molecule type: DNA
A;Residues: 1-354 <HYU>
A;Cross-references: EMBL:X59059; NID:g487722; PIDN:CAA41784.1; PID:gl334251
A;Experimental source: strain YA-335
R;Ou-Hyun, Y.; Park, Y.S.; Yum, D.Y.; Kim, J.M.; Kong, I.S.; Bai, D.H.
J. Microbiol. Biotechnol. 3, 139-145, 1993
A;Title: Nucleotide sequence and analysis of a xylanase gene (xynS) from alkali-tolerant
A;Reference number: S48126
A;Accession: S48127
A;Molecule type: DNA
A;Residues: 1-17,19-229 <JUH>
A;Cross-references: EMBL:X59059
A;Experimental source: strain YA-335

A;Note: the authors translated the initiation codon TTG for residue 1 as Leu
C;Genetics:
A;Gene: xynY
A;Start codon: TTG
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-xylosidic bonds in xylans
A;Pathway: xylan degradation
C;Superfamily: Clostridium endo-1,4-beta-xylanase B; endo-1,4-beta-xylanase homology
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-354/Product: endo-1,4-beta-xylanase Y #status predicted <MAT>
F;39-223/Domain: endo-1,4-beta-xylanase homology <XYL>
F;120,210/Active site: Glu #status predicted

Query Match 1.4%; Score 8; DB 1; Length 354;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 MTWVLGIG 60
|||||
DB 10 MTWVLGIG 17

RESULT 10
G87522
hypothetical protein CC2208 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: G87522
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.F.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon-
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87522
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-412 <STO>
A;Cross-references: GB:AE005673; NID:gl3423711; PIDN:AAK24179.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC2208

Query Match 1.4%; Score 8; DB 2; Length 412;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 LFGLLTLLI 206
|||||
DB 80 LFGLLTLLI 87

RESULT 11
C96769
unknown protein F9E11.3 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: C96769
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.
C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, G.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C96769
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-445 <STO>

A;Cross-references: GB:AE005173; NID:g10092422; PIDN:AAG12827.1; GSPDB:GN00141
 C;Genetics:
 A;Gene: F9E11.3
 A;Map position: 1

Query Match 1.4%; Score 8; DB 2; Length 445;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 SLGILLSL 100
 |||||
 Db 79 SLGILLSL 86

RESULT 12

WQBST

phosphotransferase system enzyme II (EC 2.7.1.69), sucrose-specific, factor II - Salmonella typhimurium
 N;Alternate names: phosphotransferase system enzyme II-sucrose; protein-Npi-phosphohistidine kinase
 C;Species: Salmonella typhimurium
 C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 11-Jun-1999
 C;Accession: S01036; S62329; S35016
 R;Ebner, R.; Lengeler, J.W.
 Mol. Microbiol. 2, 9-17, 1988

A;Title: DNA sequence of the gene scrA encoding the sucrose transport protein enzyme-II (scrA).
 A;Reference number: S01036; MUID:88216186; PMID:3285123
 A;Accession: S01036
 A;Molecule type: DNA
 A;Residues: 1-455.<EBN>
 A;Cross-references: EMBL:Y00541; NID:g47925; PIDN:CAA68605.1; PID:g47926
 R;Titgemeyer, F.; Jahreis, K.; Ebner, R.; Lengeler, J.W.
 Mol. Gen. Genet. 250, 197-206, 1996

A;Title: Molecular analysis of the scrA and scrB genes from Klebsiella pneumoniae and plasmid pUC19.

A;Reference number: S62329; MUID:96188840; PMID:8628219

A;Accession: S62329

A;Molecule type: DNA

A;Residues: 1-7,9-24,'HC',26-36,38-46,49-76,78-140,'N','A',145-207,'E',211-230,'V','EKL'

R;Jahreis, K.; Lengeler, J.W.

Mol. Microbiol. 9, 195-209, 1993

A;Title: Molecular analysis of two ScrR repressors and of a ScrR-FruR hybrid repressor from Klebsiella pneumoniae.

A;Reference number: S35014; MUID:94018607; PMID:8412665

A;Accession: S35016

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-24,'HC',26-435,'IAFGVAFV',445-455<JAH>

A;Cross-references: EMBL:X67750; NID:g395262; PIDN:CAA47973.1; PID:g395263

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992

C;Genetics:

A;Gene: scrA

C;Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II

C;Keywords: membrane protein; phosphoprotein; phosphotransferase; sugar transport system

F;308/Active site: His #status predicted

Query Match 1.4%; Score 8; DB 1; Length 455;

Best Local Similarity 100.0%; Pred. No. 29;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 AGWLAGLL 113
 |||||
 Db 287 AGWLAGLL 294

RESULT 13

S62331

phosphotransferase system enzyme II (EC 2.7.1.69), sucrose-specific - Klebsiella pneumoniae

C;Species: Klebsiella pneumoniae

C;Date: 08-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 18-Jun-1999

C;Accession: S62331; S15195

R;Titgemeyer, F.; Jahreis, K.; Ebner, R.; Lengeler, J.W.

Mol. Gen. Genet. 250, 197-206, 1996

A;Title: Molecular analysis of the scrA and scrB genes from Klebsiella pneumoniae and plasmid pUC19.

sucrose-6-phosphate invertase.

A;Reference number: S62329; MUID:96188840; PMID:8628219

A;Accession: S62331

A;Molecule type: DNA

A;Residues: 1-456<TIT>

A;Cross-references: EMBL:X57401; NID:g43931; PIDN:CAA40658.1; PID:g43933

R;Schmid, K.; Ebner, R.; Jahreis, K.; Lengeler, J.W.; Titgemeyer, F.

Mol. Microbiol. 5, 941-950, 1991

A;Title: A sugar-specific porin, ScrY, is involved in sucrose uptake in enteric bacteria

A;Reference number: S15193; MUID:91312133; PMID:1649946

A;Accession: S15195

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-456<SCH>

A;Cross-references: EMBL:X57401; NID:g43931; PIDN:CAA40658.1; PID:g43933

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1991

C;Genetics:

A;Gene: scrA

C;Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II

C;Keywords: membrane protein; phosphotransferase; sugar transport system

Query Match 1.4%; Score 8; DB 2; Length 456;

Best Local Similarity 100.0%; Pred. No. 29;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 AGWLAGLL 113
 |||||
 Db 288 AGWLAGLL 295

RESULT 14

A86234

hypothetical protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

A;Accession: A86234

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: A86234

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-474<STO>

A;Cross-references: GB:AE005172; NID:g2160188; PIDN:AAB60751.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

C;Superfamily: common tobacco sucrose transport protein

Query Match 1.4%; Score 8; DB 2; Length 474;

Best Local Similarity 100.0%; Pred. No. 30;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 YGRRRPFI 89
 |||||
 Db 107 YGRRRPFI 114

RESULT 15

D70600

probable polynucleotide polymerase - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C;Accession: D70600

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: D70600
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-480 <COL>
 A;Cross-references: GB:294121; GB:AL123456; NID:g3261736; PIDN:CAB08094.1; PID:e312270;
 C;Experimental source: strain H37Rv
 C;Genetics:
 A;Gene: pcnA

Query Match 1.4%; Score 8; DB 2; Length 480;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 254 GALLPRLH 261
 |||||
 Db 375 GALLPRLH 382

Search completed: December 3, 2003, 17:24:59
 Job time : 22 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 17:19:51 ; Search time 18 seconds
(without alignments)
1444.765 Million cell updates/sec

Title: US-09-593-793A-113

Perfect score: 553

Sequence: 1 MVQRLMVSRLLRHKAQLL.....AIYFATQVDFKSLAKYSA 553

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	1.6	1609	1	CTPI MYCLE
2	9	1.6	1625	1	CTPI MYCTU
3	8	1.4	21	1	BOH4 BOMVA
4	8	1.4	21	1	BOH5 BOMVA
5	8	1.4	137	1	BLP1 BOMVA
6	8	1.4	137	1	BLP2 BOMVA
7	8	1.4	223	1	UBL1 MOUSE
8	8	1.4	223	1	UBL1 RAT
9	8	1.4	456	1	PTSB KLEPN
10	8	1.4	456	1	PTSB SALTY
11	8	1.4	745	1	NLDL RAT
12	8	1.4	1158	1	ALAI ARATH
13	7	1.3	20	1	MAX7 BOMMX
14	7	1.3	78	1	EX7S PARZE
15	7	1.3	89	1	CYC6 SPIMA
16	7	1.3	128	1	VG64 BMD2
17	7	1.3	130	1	LY6E MOUSE
18	7	1.3	144	1	MAX4 BOMMX
19	7	1.3	145	1	SODM STRAI
20	7	1.3	149	1	YQCA ECOLI
21	7	1.3	151	1	YPH7 CHRVI
22	7	1.3	151	1	YQCA ERWCA
23	7	1.3	173	1	YL18 AQUAE
24	7	1.3	176	1	AROK MYCTU
25	7	1.3	182	1	RBS1 MESCR
26	7	1.3	183	1	YM39 MARPO
27	7	1.3	195	1	YBJE HAEIN
28	7	1.3	197	1	PTCA MOUSE
29	7	1.3	200	1	WINI SOLTU
30	7	1.3	206	1	PTCA HUMAN
31	7	1.3	220	1	TEGP HSVB4
32	7	1.3	226	1	COLI ONCKE
33	7	1.3	227	1	YB61 DROME

RESULT 1

ID	CTPI MYCLE	STANDARD;	PRT;	1609 AA.
AC	OS3114;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Probable cation-transporting ATPase I (EC 3.6.3.-)			
GN	CTPI OR ML2671 OR MLCB1913.02.			
OS	Mycobacterium leprae.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.			
OX	NCBI_TaxID=1769;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=TN;			
RX	MEDLINE=21128732; PubMed=11234002;			
RA	Cole S.T., Bigmeier K., Parkhill J., James K.D., Thomson N.R.,			
RA	Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,			
RA	Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,			
RA	Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,			
RA	Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,			
RA	Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,			
RA	Squires S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,			
RA	Barrell B.G.;			
RT	"Massive gene decay in the leprosy bacillus.";			
RL	Nature 409:1007-1011(2001)			
CC	- - CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.			
CC	- - SUBCELLULAR LOCATION: Integral membrane protein.			
CC	- - SIMILARITY: Belongs to the cation transport ATPases family (P-type ATPases).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; AL022118; CAA17934.1; -			
CC	EMBL; AL583926; CAC32203.1; -			
DR	PIR; E87243; E87243.			
DR	Leptoma; ML2671; -			
DR	InterPro; IPR001757; ATPase_E1-E2.			
DR	InterPro; IPR000695; H ATPase.			
DR	InterPro; IPR005834; Hydrolase.			
DR	Pfam; PR00122; E1-E2 ATPase; 1.			
DR	Pfam; PF00702; Hydrolase; 1.			
DR	PRINTS; PR00119; CATATPASE.			
DR	PRINTS; PR00120; HATPASE.			
DR	TIGRFAMs; TIGR01494; ATPase_P-type; 5.			
DR	PROSITE; PS00154; ATPase E1 E2; 1.			
KW	Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding; Complete proteome.			

Q04618 oncorhynchu
P42250 bacillus su
P46692 gallus gall
Q95336 homo sapien
Q9w777 thermotoga
Q99848 homo sapien
Q9d903 mus musculu
P32740 caenorhabdi
Q92254 squallus aca
P37763 neisseria g
Q03723 saccharomyc
P41654 methanobact

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; Z74410; CAA98940.1; -.
CC EMBL; AE006923; AAK44338.1; -.
CC PIR; C70752; C70752.
CC TIGR; MT0116; -.
CC -----
CC InterPro; IPR001757; ATPase E1-E2.
CC InterPro; IPR000695; H ATPase.
CC InterPro; IPR005834; Hydrolase.
CC Pfam; PF00122; E1-E2 ATPase; 1.
CC Pfam; PF00702; Hydrolase; 1.
CC PRINTS; PR00119; CATAPASE.
CC PRINTS; PR00120; HATPASE.
CC TIGR; TIGR01494; ATPase P-type; 6.
CC PROSITE; PS00154; ATPASE_E1_E2; 1.
CC Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;
KW Complete proteome.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 148 168 POTENTIAL.
FT TRANSMEM 177 197 POTENTIAL.
FT TRANSMEM 315 335 POTENTIAL.
FT TRANSMEM 358 378 POTENTIAL.
FT TRANSMEM 637 657 POTENTIAL.
FT TRANSMEM 673 693 POTENTIAL.
FT TRANSMEM 778 798 POTENTIAL.
FT TRANSMEM 921 941 POTENTIAL.
FT TRANSMEM 969 989 POTENTIAL.
FT TRANSMEM 997 1017 POTENTIAL.
FT TRANSMEM 1401 1421 POTENTIAL.
FT TRANSMEM 1432 1452 POTENTIAL.
FT TRANSMEM 1547 1567 POTENTIAL.
FT MOD_RES 1053 1053 PHOSPHORYLATION (BY SIMILARITY).
FT METAL 1340 1340 MAGNESIUM (BY SIMILARITY).
FT METAL 1344 1344 MAGNESIUM (BY SIMILARITY).
FT CONFLICT 1571 1625 DERDPAEDSDAGG -> SGVPRWDRSPGRASSAPRQEPPO
FT SORWHRSGWQAOQSVSCNLMNALTTTKTLTRVDRTYRRRP
FT (IN REF. 1)
SQ SEQUENCE 1625 AA; 167805 MW; 54082AD7064C22EB CRC64;

Query Match 1.6%; Score 9; DB 1; Length 1625;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QLLLVNLLT 25
Db 1430 QLLLVNLLT 1438

RESULT 3
BOH4 BOMVA
ID BOH4 BOMVA STANDARD; PRT; 21 AA.
AC P82284;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Bombinin H4.
OS Bombina variegata (Yellow-bellied toad).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
OX NCBI_TaxID=8348;
RN [1]
RP SEQUENCE, AND POST-TRANSLATIONAL MODIFICATIONS.
RC TISSUE=Skin secretion;
RX MEDLINE=94038967; PubMed=8223491;
RA Mignogna G., Simmaco M., Kreil G., Barria D.;
RT "Antibacterial and haemolytic peptides containing D-alloisoleucine

```

RT from the skin of Bombina variegata.";
 RL EMO J. 12:4829-4832(1993).
 CC -1- FUNCTION: HAS ANTIMICROBIAL AND HEMOLYTIC ACTIVITIES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Skin.
 CC -1- SIMILARITY: BELONGS TO THE BOMBININ FAMILY.
 KW Amphibian defense peptide; Antibiotic; Hemolysis; Amidation;
 KW D-amino acid.
 FT PEPTIDE 1 20 BOMBININ H4.
 FT MOD RES 2 2 D-ALLO-ISOLEUCINE.
 FT MOD RES 20 20 AMIDATION (G-21 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 21 AA; 1975 MW; 08C7281E1D6D43BD CRC64;
 Query Match 1.4%; Score 8; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.98;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 59 IGPVLGLV 66
 Db 2 IGPVLGLV 9
 RESULT 4
 BOH5_BOMVA STANDARD; PRT; 21 AA.
 AC P82285;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Bombinin H5.
 OS Bombina variegata (Yellow-bellied toad).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
 OX NCBI_TaxID=8348;
 RN [1]
 RP SEQUENCE, AND POST-TRANSLATIONAL MODIFICATIONS.
 RC TISSUE=Skin secretion;
 RA Mignogna G., Simmaco M., Kreil G., Barra D.;
 RT "Antibacterial and haemolytic peptides containing D-alloisoleucine
 from the skin of Bombina variegata.";
 RL EMO J. 12:4829-4832(1993).
 CC -1- FUNCTION: HAS ANTIMICROBIAL AND HEMOLYTIC ACTIVITIES.
 CC -1- SIMILARITY: BELONGS TO THE BOMBININ FAMILY.
 KW Amphibian defense peptide; Antibiotic; Hemolysis; Amidation;
 KW D-amino acid.
 FT PEPTIDE 1 20 BOMBININ H5.
 FT MOD RES 2 2 D-ALLO-ISOLEUCINE.
 FT MOD RES 20 20 AMIDATION (G-21 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 21 AA; 1975 MW; 08C7281E1D6D43BD CRC64;
 Query Match 1.4%; Score 8; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.98;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 59 IGPVLGLV 66
 Db 2 IGPVLGLV 9

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
 OX NCBI_TaxID=8348;
 RN [1]
 RP SEQUENCE FROM N.A., AND PRELIMINARY SEQUENCE OF 44-69.
 RC TISSUE=Skin, and Skin secretion;
 RX MEDLINE=91293126; PubMed=1712299;
 RA Simmaco M., Barra D., Chiarini F., Novello L., Melchiorri P.,
 RA Kreil G., Richter K.;
 RT "A family of bombinin-related peptides from the skin of Bombina
 variegata";
 RL Eur. J. Biochem. 199:217-222(1991).
 RN [2]
 RP SEQUENCE OF 117-136, AND POST-TRANSLATIONAL MODIFICATIONS.
 RC TISSUE=Skin secretion;
 RX MEDLINE=94038967; PubMed=8223491;
 RA Mignogna G., Simmaco M., Kreil G., Barra D.;
 RT "Antibacterial and haemolytic peptides containing D-alloisoleucine
 from the skin of Bombina variegata.";
 RL EMO J. 12:4829-4832(1993).
 CC -1- FUNCTION: HAS ANTIMICROBIAL ACTIVITY, BUT NO HEMOLYTIC ACTIVITY.
 CC PRELIMINARY EVIDENCE INDICATES THAT THIS PEPTIDE DOES NOT LYSE AND
 THUS KILL THE BACTERIA BY ITS ANTIMICROBIAL ACTIVITY.
 CC -1- FUNCTION: BOMBININ H HAS ANTIBACTERIAL AND HEMOLYTIC ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Skin.
 CC -1- SIMILARITY: BELONGS TO THE BOMBININ FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; X59695; CAA42216.1;
 DR PIR; S16222; S16222.
 DR Pfam; PF05298; Bombinin; 1.
 KW Amphibian defense peptide; Antibiotic; Amidation;
 KW Cleavage on pair of basic residues; Signal; Hemolysis; D-amino acid.
 FT SIGNAL 1 18 POTENTIAL.
 FT PEPTIDE 19 43 ACIDIC PEPTIDE 1-1 (POTENTIAL).
 FT PEPTIDE 44 70 BOMBININ-LIKE PEPTIDE 1.
 FT PEPTIDE 74 81 OCTAPEPTIDE (POTENTIAL).
 FT PEPTIDE 84 114 ACIDIC PEPTIDE 1-2 (POTENTIAL).
 FT PEPTIDE 117 136 BOMBININ H.
 FT MOD RES 70 70 AMIDATION (G-71 PROVIDE AMIDE GROUP).
 FT MOD RES 118 118 D-ALLO-ISOLEUCINE.
 FT MOD RES 136 136 AMIDATION (G-137 PROVIDE AMIDE GROUP).
 FT VARIANT 117 117 I -> L.
 FT VARIANT 124 124 L -> M.
 SQ SEQUENCE 137 AA; 14982 MW; 3EC3EF6E47CA1F92 CRC64;
 Query Match 1.4%; Score 8; DB 1; Length 137;
 Best Local Similarity 100.0%; Pred. No. 4.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 59 IGPVLGLV 66
 Db 118 IGPVLGLV 125
 RESULT 6
 BLP2_BOMVA STANDARD; PRT; 137 AA.
 ID BLP2_BOMVA
 AC P82286;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Bombinin-like peptides 2 precursor [Contains: Acidic peptide 2-1;
 DE Bombinin-like peptide 2 (BLP-2); Octapeptide 2; Acidic peptide 2-2;
 DE Bombinin H2].

OS Bombina variegata (Yellow-bellied toad).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
OX NCBI_TaxID=8348;
[1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 117-136.
RP TISSUE=Skin, and Skin secretion;
RX MEDLINE=99266278; PubMed=10333736;
RA Simmaco M., Mignogna G., Barra D.;
RT "Antimicrobial peptides from amphibian skin: what do they tell us?";
RL Biopolymers 47:435-450(1998).
[2]
RN SEQUENCE OF 117-136, AND POST-TRANSLATIONAL MODIFICATIONS.
RP TISSUE=Skin secretion;
RX MEDLINE=94038967; PubMed=8223491;
RA Mignogna G., Simmaco M., Kreil G., Barra D.;
RT "Antibacterial and haemolytic peptides containing D-alloisoleucine
from the skin of Bombina variegata.";
RL EMBO J. 12:4829-4832(1993).
CC -!- FUNCTION: BOMBININ-LIKE PEPTIDE 2 HAS ANTIMICROBIAL ACTIVITY, BUT
NO HEMOLYTIC ACTIVITY. PRELIMINARY EVIDENCE INDICATES THAT THIS
PEPTIDE DOES NOT LYSIS AND THUS KILL THE BACTERIA BY ITS
ANTIMICROBIAL ACTIVITY.
CC -!- FUNCTION: BOMBININ H2 HAS ANTIBACTERIAL AND HEMOLYTIC ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- SIMILARITY: BELONGS TO THE BOMBININ FAMILY.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AJ251565; CAB61443.1; --
DR Pfam; PF05298; Bombinin; 1.
KW Amphibian defense peptide; Antibiotic; Amidation;
KW Cleavage on pair of basic residues; Signal; Hemolysis.
FT SIGNAL 1 18
FT PEPTIDE 19 43 POTENTIAL.
FT PEPTIDE 44 70 BOMBININ-LIKE PEPTIDE 2.
FT PEPTIDE 74 81 OCTAPEPTIDE (POTENTIAL).
FT PEPTIDE 84 114 ACIDIC PEPTIDE 2-2 (POTENTIAL).
FT PEPTIDE 117 136 BOMBININ H2.
FT MOD RES 70 70 AMIDATION (G-71 PROVIDE AMIDE GROUP).
FT MOD RES 136 136 AMIDATION (G-137 PROVIDE AMIDE GROUP).
SQ SEQUENCE 137 AA; 15035 MW; 6BFBD1A5F0B6CCD6 CRC64;

Query Match 1.4%; Score 8; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 IGPVLGLV 66
Db 118 IGPVLGLV 125
|||||||
|

RESULT 7
UB11 MOUSE
ID UB11 MOUSE STANDARD; PRT; 223 AA.
AC Q9R0F9; Q9R122;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase isozyme L1 (EC 3.4.19.12) (UCH-
DE L1) (Ubiquitin thiolesterase L1) (Neuron cytoplasmic protein 9.5)
DE (RGP 9.5) (RGP9.5).
GN UCHL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99400551; PubMed=10471497;
RA Saigoh K., Wang Y.-L., Suh J.G., Yamanishi T., Sakai Y., Kiyosawa H.,
RA Harada T., Ichihara N., Wakana S., Kikuchi T., Wada K.;
RT "Intragenic deletion in the gene encoding ubiquitin carboxy-terminal
hydrolase in gad mice.";
RL Nat. Genet. 23:47-51(1999).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Pituitary;
RA Marzban G., Grillari J., Reisinger E., Hemetseberger T.,
RA Hohenwarter O., Katinger H.;
RT "Cloning of the mouse homologue of ubiquitin carboxyl-terminal
hydrolase PGP9.5.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Hippocampus;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gless C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- FUNCTION: UBIQUITIN-PROTEIN HYDROLASE IS INVOLVED BOTH IN THE
PROCESSING OF UBIQUITIN PRECURSORS AND OF UBIQUINATED PROTEINS.
CC THIS ENZYME IS A THIOL PROTEASE THAT RECOGNIZE AND HYDROLYZE
A PEPTIDE BOND AT THE C-TERMINAL GLYCINE OF UBIQUITIN.
CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
ubiquitin + a thiol.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in brain and testis.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C12.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AB025313; BAA84083.1; --
DR EMBL; AF172334; RAD51029.1; --
DR EMBL; AK013729; BAB28976.1; --
DR HSSP; F15374; UCH.
DR MEROPS; C12.001; --
DR MGD; MGI:103149; Uchl1.
DR InterPro; IPR001578; UCH 1.
DR Pfam; PF01088; Peptidase_C12; 1.
DR PRINTS; PR00707; UBCHYDLASE.
DR PRODOM; PD350662; UCH 1; 1.
DR PROSITE; PS00140; UCH 1; 1.
KW Ub1 conjugation pathway; Hydrolase; Thiol protease; Multigene family.
FT ACT_SITE 90 90 BY SIMILARITY.

FT TRANSMEM 181 201 POTENTIAL.
 FT TRANSMEM 209 229 POTENTIAL.
 FT TRANSMEM 247 267 POTENTIAL.
 FT TRANSMEM 288 308 POTENTIAL.
 FT TRANSMEM 330 350 POTENTIAL.
 FT TRANSMEM 360 380 POTENTIAL.
 FT TRANSMEM 388 408 POTENTIAL.
 FT TRANSMEM 428 448 POTENTIAL.
 FT MOD_RES 25 26 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 309 309 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 456 AA; 48022 MW; 4AFDF5405CAEPC66 CRC64;

Query Match 1.4%; Score 8; DB 1; Length 456;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 AGWLAGLL 113
 Db 288 AGWLAGLL 295
 |||||

RESULT 10
 PTSSB_SALTY STANDARD; PRT; 456 AA.
 AC P08470;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-JUN-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE PTS system, sucrose-specific IIBC component (EIIBC-SCR) (Sucrose-
 DE permease IIBC component) (Phosphotransferase enzyme II, BC component)
 DE [EC 2.7.1.69] (EIIB-SCR).
 GN SCRA.
 OS Salmonella typhimurium.
 OG Plasmid pUR400.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88216186; PubMed=3285123;
 RA Ebner R., Lengeler J.W.;
 RT "DNA sequence of the gene scrA encoding the sucrose transport protein
 RT EnzymeII(Scr) of the phosphotransferase system from enteric bacteria:
 RT homology of the EnzymeII(Scr) and EnzymeII(Bgl) proteins.";
 RL Mol. Microbiol. 2:9-17(1988).
 RN [2]
 RP REVISIONS.
 RX MEDLINE=94018607; PubMed=8412665;
 RA Jahreis K., Lengeler J.W.;
 RT "Molecular analysis of two ScrR repressors and of a ScrR-FruR hybrid
 RT repressor for sucrose and D-fructose specific regulons from enteric
 RT bacteria.";
 RL Mol. Microbiol. 9:195-209(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96188840; PubMed=8628219;
 RA Titgemeyer F., Jahreis K., Ebner R., Lengeler J.W.;
 RT "Molecular analysis of the scrA and scrB genes from Klebsiella
 RT pneumoniae and plasmid pUR400, which encode the sucrose transport
 RT protein Enzyme II Scr of the phosphotransferase system and a
 RT sucrose-6-phosphate invertase.";
 RL Mol. Gen. Genet. 250:197-206(1996).
 RN [4]
 RP SEQUENCE OF 1-7 FROM N.A.
 RC STRAIN=6153-62;
 RX MEDLINE=91100329; PubMed=1846143;
 RA Hardesty C., Ferran C., Drenzo J.M.;
 RT "Plasmid-mediated sucrose metabolism in Escherichia coli:
 RT characterization of scrI, the structural gene for a
 RT phosphoenolpyruvate-dependent sucrose phosphotransferase system
 RT outer membrane porin.";
 RL J. Bacteriol. 173:449-456(1991).
 CC -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT

SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
 -TRANSPORT SYSTEM. THE IID DOMAINS CONTAIN THE SUGAR BINDING SITE
 AND THE TRANSMEMBRANE CHANNEL; THE IIA DOMAIN CONTAINS THE PRIMARY
 PHOSPHORYLATION SITE (THE DONOR IS PHOSHO-HPR); IIA TRANSFERS ITS
 PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
 THE SUGAR. EIIBC-SCR BELONGS TO THE EIIBC DOMAIN TYPE AND LACKS A
 COVALENTLY BOUND IIA DOMAIN. INSTEAD, EIIB-SCR-MEDIATED
 PHOSPHORYLATION OF SUCROSE REQUIRES THE ACTIVITY OF ENZYME IIA-
 GLC, A COMPONENT OF THE MAJOR GLUCOSE TRANSPORT SYSTEM.
 -!- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
 histidine + sugar phosphate.
 -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 -!- SIMILARITY: Contains 1 PTS EIIB domain.
 -!- SIMILARITY: Contains 1 PTS EIIC domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch)

 EMBL; Y00541; CAA68605.1; ALT_SEQ.
 DR EMBL; X67750; CAA49793.1; -;
 DR EMBL; M38416; AAA98418.1; -;
 DR PIR; S01036; WOEST.
 DR InterPro; IPR001996; PTS_EIIB.
 DR InterPro; IPR003352; PTS_EIIC.
 DR InterPro; IPR004719; PTSIIC_glc.
 DR Pfam; PF00367; PTS_EIIB; 1.
 DR Pfam; PF02378; PTS_EIIC; 1.
 DR ProDom; PD001476; PTS_EIIB; 1.
 DR TIGRFAMs; TIGR00826; EIIB_glc; 1.
 DR TIGRFAMs; TIGR00852; pts_glc; 1.
 DR PROSITE; PS01035; PTS_EIIB_CYS; 1.
 KW Phosphotransferase system; Transferrase; Transport; Sugar transport;
 KW Plasmid; Transmembrane; Inner membrane; Phosphorylation.
 FT DOMAIN 1 111 EIIB DOMAIN.
 FT DOMAIN 112 456 EIIC DOMAIN.
 FT TRANSMEM 144 164 POTENTIAL.
 FT TRANSMEM 181 201 POTENTIAL.
 FT TRANSMEM 213 233 POTENTIAL.
 FT TRANSMEM 247 267 POTENTIAL.
 FT TRANSMEM 288 308 POTENTIAL.
 FT TRANSMEM 330 350 POTENTIAL.
 FT TRANSMEM 360 380 POTENTIAL.
 FT TRANSMEM 388 408 POTENTIAL.
 FT TRANSMEM 428 448 POTENTIAL.
 FT MOD_RES 26 26 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 309 309 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 456 AA; 47864 MW; 98A6F1620AE50885 CRC64;

Query Match 1.4%; Score 8; DB 1; Length 456;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 AGWLAGLL 113
 Db 288 AGWLAGLL 295
 |||||

RESULT 11
 NLDL RAT STANDARD; PRT; 745 AA.
 AC O54637;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE N-acetylated-alpha-linked acidic dipeptidase like protein
 DE (EC 3.4.17.21) (NAALADase L) (Ileal dipeptidylpeptidase) (100 kDa
 DE ileum brush border membrane protein) (I100).


```

FT DOMAIN          935 948 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM        949 968 POTENTIAL.
FT DOMAIN          969 998 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM        999 1020 POTENTIAL.
FT DOMAIN         1021 1027 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM        1028 1050 POTENTIAL.
FT DOMAIN         1051 1056 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM        1057 1077 POTENTIAL.
FT DOMAIN         1078 1090 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM        1091 1115 POTENTIAL.
FT DOMAIN         1116 1158 CYTOPLASMIC (POTENTIAL).
FT MOD RES         457 457 PHOSPHORYLATION (BY SIMILARITY).
FT METAL           859 859 MAGNESIUM (BY SIMILARITY).
FT METAL           863 863 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE        1158 AA; 130328 MW; 5CC042B40C8C974D CRC64;

Query Match          1.4%; Score 8; DB 1; Length 1158;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 ILLSLFLI 103
DB 331 ILLSLFLI 338

RESULT 13
MAX7_BOMMX
ID MAX7_BOMMX STANDARD; PRT; 20 AA.
AC P83086;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Maximin 7.
OS Bombina maxima (Giant fire-bellied toad) (Chinese red belly toad).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
OX NCBI_TaxID=161274;
RN [1]
RP SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC TISSUE=Skin secretion;
RA Chen T.B., McClean S., Orr D.F., Bjourson A.J., Rao P.F., Shaw C.;
RT "Isolation and structural characterisation of antimicrobial peptides
RT from the venom of the Chinese large-webbed bell toad (Bombina
RT maxima).";
RL Submitted (JUL-2001) to the SWISS-PROT data bank.
CC -!- FUNCTION: Has antimicrobial activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Antibiotic.
SQ SEQUENCE 20 AA; 1947 MW; EC7541AA1BF3BD49 CRC64;

Query Match          1.3%; Score 7; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 GPVGLV 56
DB 3 GPVGLV 9

RESULT 14
EX7S_PARZE
ID EX7S_PARZE STANDARD; PRT; 78 AA.
AC Q8LIH9;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
DE (Exonuclease VII small subunit).
GN XSEB.
OS Paracoccus zeaxanthinifaciens.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;

Rhodobacteraceae; Paracoccus.
OX NCBI_TaxID=187400;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R114;
RA Huemelin M.;
RT "Genetics of isoprenoid biosynthesis in Paracoccus
RT zeaxanthinifaciens.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE XSEB FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ431697; CAD24424.1; --
DR HAMAP; MF_00337; --; 1.
DR InterPro; IPR003761; Exonuc VII S.
DR Pfam; PF02609; Exonuc VII S; 1.
DR TIGRFAMs; TIGR01280; xseB; 1.
KW Hydrolase; Nuclease; Exonuclease.
SQ SEQUENCE 78 AA; 8492 MW; 1D3963A65D5C53D8 CRC64;

Query Match          1.3%; Score 7; DB 1; Length 78;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 226 TEPAEGL 232
DB 72 TEPAEGL 78

RESULT 15
CYC6_SPIMA
ID CYC6_SPIMA STANDARD; PRT; 89 AA.
AC P00118;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome c6 (Soluble cytochrome f) (Cytochrome c553) (Cytochrome c-
DE 553).
GN PETJ.
OS Spirulina maxima (Arthrospira maxima).
OC Bacteria; Cyanobacteria; Oscillatoriales; Spirulina.
OX NCBI_TaxID=1155;
RN [1]
RP SEQUENCE.
RX MEDLINE=75100362; PubMed=803642;
RA Ambler R.P., Bartsch R.G.;
RT "Amino acid sequence similarity between cytochrome f from a
RT blue-green bacterium and algal chloroplasts.";
RL Nature 253:285-288 (1975).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX MEDLINE=21371781; PubMed=11478889;
RA Sawaya M.R., Krogmann D.W., Serag A., Ho K.K., Yeates T.O.,
RA Kerfeld C.A.;
RT "Structures of cytochrome c-549 and cytochrome c6 from the
RT cyanobacterium Arthrospira maxima."
RL Biochemistry 40:9215-9225 (2001).
RN [3]

```

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).
RX MEDLINE=22071407; PubMed=12077429;
RA Kerfeld C.A.; Sawaya M.R.; Krogmann D.W.; Yeates T.O.;
RT "Structure of cytochrome c6 from Arthrospira maxima: an assembly of 24
RT subunits in a nearly symmetric shell".
RL Acta Crystallogr. D 58:1104-1110(2002).
CC -!- FUNCTION: Functions as an electron carrier between membrane-bound
CC cytochrome b6f and photosystem I in oxygenic photosynthesis.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- PTM: Binds one heme group per molecule.
CC -!- MISCELLANEOUS: The midpoint redox potential for this protein is
CC +314 mV.
CC -!- SIMILARITY: Belongs to the cytochrome c family. PctJ subfamily.
DR PIR; A00110; CCSG6.
DR PDB; 1F1F; 08-AUG-01.
DR PDB; 1K1B; 03-JUL-02.
DR HAMAP; MF_00594; -; 1.
DR InterPro; IPR003088; Cyt_C1.
DR InterPro; IPR002329; Cyt_C1C.
DR DR Pfam; PF00034; cytochrome_c; 1.
DR PRINTS; PR00605; CYTOCHROME_C.
DR PROSITE; PS00130; CYTOCHROME_C; 1.
KW Electron transport; Photosynthesis; Heme; 3D-structure.
FT BINDING 14 14 HEME (COVALENT).
FT BINDING 17 17 HEME (COVALENT).
FT METAL 18 18 IRON (HEME AXIAL LIGAND).
FT METAL 62 62 IRON (HEME AXIAL LIGAND).
FT HELIX 3 13
FT TURN 14 14
FT HELIX 15 18
FT HELIX 19 21
FT TURN 27 28
FT HELIX 33 39
FT TURN 41 45
FT HELIX 47 57
FT STRAND 59 59
FT TURN 60 61
FT STRAND 62 62
FT TURN 66 68
FT HELIX 71 87
FT TURN 88 89
SQ SEQUENCE 89 AA; 9236 MW; 4796D56B3E8A9AF85 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 545 KSDLAKY 551
Db 33 KSDLAKY 39
|||||

Search completed: December 3, 2003, 17:23:33
Job time : 20 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)